

sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such PRO polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the PRO polypeptide fragments encoded by these nucleotide molecule fragments, preferably those PRO polypeptide fragments that comprise a binding site for an anti-PRO antibody.

5 In another embodiment, the invention provides isolated PRO polypeptide encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a certain aspect, the invention concerns an isolated PRO polypeptide, comprising an amino acid sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet
10 more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity, yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about
15 about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to a PRO polypeptide having a full-length amino acid sequence as disclosed herein, a full-length amino acid sequence lacking the signal peptide as disclosed herein or an
20 extracellular domain of a transmembrane protein as disclosed herein.

In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence having at least about 80% sequence identity, preferably at least about 81% sequence identity, more preferably at least about 82% sequence identity, yet more preferably at least about 83% sequence identity, yet more preferably at least about 84% sequence identity, yet more preferably at least about 85% sequence identity,
25 yet more preferably at least about 86% sequence identity, yet more preferably at least about 87% sequence identity, yet more preferably at least about 88% sequence identity, yet more preferably at least about 89% sequence identity, yet more preferably at least about 90% sequence identity, yet more preferably at least about 91% sequence identity, yet more preferably at least about 92% sequence identity, yet more preferably at least about 93% sequence identity, yet more preferably at least about 94% sequence identity, yet more preferably at least about 95% sequence identity, yet more preferably at least about 96% sequence identity, yet more preferably at least about 97% sequence identity, yet more preferably at least about 98% sequence identity and yet more preferably at least about 99% sequence identity to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.
30

In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid
35 sequence scoring at least about 80% positives, preferably at least about 81% positives, more preferably at least about 82% positives, yet more preferably at least about 83% positives, yet more preferably at least about 84% positives, yet more preferably at least about 85% positives, yet more preferably at least about 86% positives,

yet more preferably at least about 87% positives, yet more preferably at least about 88% positives, yet more preferably at least about 89% positives, yet more preferably at least about 90% positives, yet more preferably at least about 91% positives, yet more preferably at least about 92% positives, yet more preferably at least about 93% positives, yet more preferably at least about 94% positives, yet more preferably at least about 95% positives, yet more preferably at least about 96% positives, yet more preferably at least about 97% positives, yet more preferably at least about 98% positives and yet more preferably at least about 99% positives when compared with the amino acid sequence of a PRO polypeptide having a full-length amino acid sequence as disclosed herein, a full-length amino acid sequence lacking the signal peptide as disclosed herein or an extracellular domain of a transmembrane protein as disclosed herein.

In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

Another aspect the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist or antagonist thereof as hereinbefore described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist or antagonist thereof or an anti-PRO antibody.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a native sequence PRO281 (UNQ244) cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA16422-1209".

Figure 2 shows the amino acid sequence (SEQ ID NO:2) derived from the coding sequence of SEQ ID NO:1 shown in Figure 1.

Figure 3 shows a nucleotide sequence (SEQ ID NO:5) of a native sequence PRO276 (UNQ243) cDNA, wherein SEQ ID NO:5 is a clone designated herein as "DNA16435-1208".

Figure 4 shows the amino acid sequence (SEQ ID NO:6) derived from the coding sequence of SEQ ID NO:5 shown in Figure 3.

Figure 5 shows a nucleotide sequence (SEQ ID NO:7) of a native sequence PRO189 (UNQ163) cDNA, wherein SEQ ID NO:7 is a clone designated herein as "DNA21642-1391".

Figure 6 shows the amino acid sequence (SEQ ID NO:8) derived from the coding sequence of SEQ ID NO:7 shown in Figure 5.

Figure 7 shows a nucleotide sequence designated herein as DNA14187 (SEQ ID NO:9).

Figure 8 shows a nucleotide sequence (SEQ ID NO:13) of a native sequence PRO190 (UNQ164) cDNA, wherein SEQ ID NO:13 is a clone designated herein as "DNA23334-1392".

Figure 9 shows the amino acid sequence (SEQ ID NO:14) derived from the coding sequence of SEQ ID NO:13 shown in Figure 8.

5 Figure 10 shows a nucleotide sequence designated herein as DNA14232 (SEQ ID NO:15).

Figure 11 shows a nucleotide sequence (SEQ ID NO:19) of a native sequence PRO341 (UNQ300) cDNA, wherein SEQ ID NO:19 is a clone designated herein as "DNA26288-1239".

Figure 12 shows the amino acid sequence (SEQ ID NO:20) derived from the coding sequence of SEQ ID NO:19 shown in Figure 11.

10 Figure 13 shows a nucleotide sequence designated herein as DNA12920 (SEQ ID NO:21).

Figure 14 shows a nucleotide sequence (SEQ ID NO:22) of a native sequence PRO180 (UNQ154) cDNA, wherein SEQ ID NO:22 is a clone designated herein as "DNA26843-1389".

Figure 15 shows the amino acid sequence (SEQ ID NO:23) derived from the coding sequence of SEQ ID NO:22 shown in Figure 14.

15 Figure 16 shows a nucleotide sequence designated herein as DNA12922 (SEQ ID NO:24).

Figure 17 shows a nucleotide sequence (SEQ ID NO:27) of a native sequence PRO194 (UNQ168) cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA26844-1394".

Figure 18 shows the amino acid sequence (SEQ ID NO:28) derived from the coding sequence of SEQ ID NO:27 shown in Figure 17.

20 Figure 19 shows a nucleotide sequence (SEQ ID NO:29) of a native sequence PRO203 (UNQ177) cDNA, wherein SEQ ID NO:29 is a clone designated herein as "DNA30862-1396".

Figure 20 shows the amino acid sequence (SEQ ID NO:30) derived from the coding sequence of SEQ ID NO:29 shown in Figure 19.

Figure 21 shows a nucleotide sequence designated herein as DNA15618 (SEQ ID NO:31).

25 Figure 22 shows a nucleotide sequence (SEQ ID NO:32) of a native sequence PRO290 (UNQ253) cDNA, wherein SEQ ID NO:32 is a clone designated herein as "DNA35680-1212".

Figure 23 shows the amino acid sequence (SEQ ID NO:33) derived from the coding sequence of SEQ ID NO:32 shown in Figure 22.

30 Figure 24 shows a nucleotide sequence (SEQ ID NO:35) of a native sequence PRO874 (UNQ441) cDNA, wherein SEQ ID NO:35 is a clone designated herein as "DNA40621-1440".

Figure 25 shows the amino acid sequence (SEQ ID NO:36) derived from the coding sequence of SEQ ID NO:35 shown in Figure 24.

Figure 26 shows a nucleotide sequence (SEQ ID NO:40) of a native sequence PRO710 (UNQ374) cDNA, wherein SEQ ID NO:40 is a clone designated herein as "DNA44161-1434".

35 Figure 27 shows the amino acid sequence (SEQ ID NO:41) derived from the coding sequence of SEQ ID NO:40 shown in Figure 26.

Figure 28 shows a nucleotide sequence designated herein as DNA38190 (SEQ ID NO:42).

Figure 29 shows a nucleotide sequence (SEQ ID NO:46) of a native sequence PRO1151 (UNQ581) cDNA, wherein SEQ ID NO:46 is a clone designated herein as "DNA44694-1500".

Figure 30 shows the amino acid sequence (SEQ ID NO:47) derived from the coding sequence of SEQ ID NO:46 shown in Figure 29.

Figure 31 shows a nucleotide sequence (SEQ ID NO:51) of a native sequence PRO1282 (UNQ652) cDNA, wherein SEQ ID NO:51 is a clone designated herein as "DNA45495-1550".

Figure 32 shows the amino acid sequence (SEQ ID NO:52) derived from the coding sequence of SEQ ID NO:51 shown in Figure 31.

Figure 33 shows a nucleotide sequence (SEQ ID NO:56) of a native sequence PRO358 cDNA, wherein SEQ ID NO:56 is a clone designated herein as "DNA47361-1154".

Figure 34 shows the amino acid sequence (SEQ ID NO:57) derived from the coding sequence of SEQ ID NO:56 shown in Figure 33.

Figures 35A-B show a nucleotide sequence (SEQ ID NO:61) of a native sequence PRO1310 cDNA, wherein SEQ ID NO:61 is a clone designated herein as "DNA47394-1572".

Figure 36 shows the amino acid sequence (SEQ ID NO:62) derived from the coding sequence of SEQ ID NO:61 shown in Figures 35A-B.

Figure 37 shows a nucleotide sequence (SEQ ID NO:66) of a native sequence PRO698 (UNQ362) cDNA, wherein SEQ ID NO:66 is a clone designated herein as "DNA48320-1433".

Figure 38 shows the amino acid sequence (SEQ ID NO:67) derived from the coding sequence of SEQ ID NO:66 shown in Figure 37.

Figure 39 shows a nucleotide sequence designated herein as DNA39906 (SEQ ID NO:68).

Figure 40 shows a nucleotide sequence (SEQ ID NO:72) of a native sequence PRO732 (UNQ396) cDNA, wherein SEQ ID NO:72 is a clone designated herein as "DNA48334-1435".

Figure 41 shows the amino acid sequence (SEQ ID NO:73) derived from the coding sequence of SEQ ID NO:72 shown in Figure 40.

Figure 42 shows a nucleotide sequence designated herein as DNA20239 (SEQ ID NO:74).

Figure 43 shows a nucleotide sequence designated herein as DNA38050 (SEQ ID NO:75).

Figure 44 shows a nucleotide sequence designated herein as DNA40683 (SEQ ID NO:76).

Figure 45 shows a nucleotide sequence designated herein as DNA42580 (SEQ ID NO:77).

Figures 46A-B show a nucleotide sequence (SEQ ID NO:83) of a native sequence PRO1120 (UNQ559) cDNA, wherein SEQ ID NO:83 is a clone designated herein as "DNA48606-1479".

Figure 47 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:83 shown in Figures 46A-B.

Figure 48 shows a nucleotide sequence (SEQ ID NO:94) of a native sequence PRO537 (UNQ338) cDNA, wherein SEQ ID NO:94 is a clone designated herein as "DNA49141-1431".

Figure 49 shows the amino acid sequence (SEQ ID NO:95) derived from the coding sequence of SEQ ID NO:94 shown in Figure 48.

Figure 50 shows a nucleotide sequence (SEQ ID NO:96) of a native sequence PRO536 (UNQ337)

cDNA, wherein SEQ ID NO:96 is a clone designated herein as "DNA49142-1430".

Figure 51 shows the amino acid sequence (SEQ ID NO:97) derived from the coding sequence of SEQ ID NO:96 shown in Figure 50.

Figure 52 shows a nucleotide sequence (SEQ ID NO:98) of a native sequence PRO535 (UNQ336) cDNA, wherein SEQ ID NO:98 is a clone designated herein as "DNA49143-1429".

5 Figure 53 shows the amino acid sequence (SEQ ID NO:99) derived from the coding sequence of SEQ ID NO:98 shown in Figure 52.

Figure 54 shows a nucleotide sequence designated herein as DNA30861 (SEQ ID NO:100).

Figure 55 shows a nucleotide sequence designated herein as DNA36351 (SEQ ID NO:101).

10 Figure 56 shows a nucleotide sequence (SEQ ID NO:102) of a native sequence PRO718 (UNQ386) cDNA, wherein SEQ ID NO:102 is a clone designated herein as "DNA49647-1398".

Figure 57 shows the amino acid sequence (SEQ ID NO:103) derived from the coding sequence of SEQ ID NO:102 shown in Figure 56.

Figure 58 shows a nucleotide sequence designated herein as DNA15386 (SEQ ID NO:104).

Figure 59 shows a nucleotide sequence designated herein as DNA16630 (SEQ ID NO:105).

15 Figure 60 shows a nucleotide sequence designated herein as DNA16829 (SEQ ID NO:106).

Figure 61 shows a nucleotide sequence designated herein as DNA28357 (SEQ ID NO:107).

Figure 62 shows a nucleotide sequence designated herein as DNA43512 (SEQ ID NO:108).

Figure 63 shows a nucleotide sequence (SEQ ID NO:112) of a native sequence PRO872 (UNQ439) cDNA, wherein SEQ ID NO:112 is a clone designated herein as "DNA49819-1439".

20 Figure 64 shows the amino acid sequence (SEQ ID NO:113) derived from the coding sequence of SEQ ID NO:112 shown in Figure 63.

Figure 65 shows a nucleotide sequence (SEQ ID NO:114) of a native sequence PRO1063 (UNQ128) cDNA, wherein SEQ ID NO:114 is a clone designated herein as "DNA49820-1427".

25 Figure 66 shows the amino acid sequence (SEQ ID NO:115) derived from the coding sequence of SEQ ID NO:114 shown in Figure 65.

Figure 67 shows a nucleotide sequence (SEQ ID NO:116) of a native sequence PRO619 (UNQ355) cDNA, wherein SEQ ID NO:116 is a clone designated herein as "DNA49821-1562".

Figure 68 shows the amino acid sequence (SEQ ID NO:117) derived from the coding sequence of SEQ ID NO:116 shown in Figure 67.

30 Figure 69 shows a nucleotide sequence (SEQ ID NO:118) of a native sequence PRO943 (UNQ480) cDNA, wherein SEQ ID NO:118 is a clone designated herein as "DNA52192-1369".

Figure 70 shows the amino acid sequence (SEQ ID NO:119) derived from the coding sequence of SEQ ID NO:118 shown in Figure 69.

35 Figure 71 shows a nucleotide sequence (SEQ ID NO:123) of a native sequence PRO1188 (UNQ602) cDNA, wherein SEQ ID NO:123 is a clone designated herein as "DNA52598-1518".

Figure 72 shows the amino acid sequence (SEQ ID NO:124) derived from the coding sequence of SEQ ID NO:123 shown in Figure 71.

Figure 73 shows a nucleotide sequence (SEQ ID NO:128) of a native sequence PRO1133 (UNQ571) cDNA, wherein SEQ ID NO:128 is a clone designated herein as "DNA53913-1490".

Figure 74 shows the amino acid sequence (SEQ ID NO:129) derived from the coding sequence of SEQ ID NO:128 shown in Figure 73.

Figure 75 shows a nucleotide sequence (SEQ ID NO:134) of a native sequence PRO784 (UNQ459) cDNA, wherein SEQ ID NO:134 is a clone designated herein as "DNA53978-1443".

Figure 76 shows the amino acid sequence (SEQ ID NO:135) derived from the coding sequence of SEQ ID NO:134 shown in Figure 75.

Figure 77 shows a nucleotide sequence designated herein as DNA44661 (SEQ ID NO:136).

Figure 78 shows a nucleotide sequence (SEQ ID NO:137) of a native sequence PRO783 (UNQ458) cDNA, wherein SEQ ID NO:137 is a clone designated herein as "DNA53996-1442".

Figure 79 shows the amino acid sequence (SEQ ID NO:138) derived from the coding sequence of SEQ ID NO:137 shown in Figure 78.

Figure 80 shows a nucleotide sequence designated herein as DNA45201 (SEQ ID NO:139).

Figure 81 shows a nucleotide sequence designated herein as DNA14575 (SEQ ID NO:140).

Figure 82 shows a nucleotide sequence (SEQ ID NO:145) of a native sequence PRO820 (UNQ503) cDNA, wherein SEQ ID NO:145 is a clone designated herein as "DNA56041-1416".

Figure 83 shows the amino acid sequence (SEQ ID NO:146) derived from the coding sequence of SEQ ID NO:145 shown in Figure 82.

Figure 84 shows a nucleotide sequence (SEQ ID NO:147) of a native sequence PRO1080 (UNQ537) cDNA, wherein SEQ ID NO:147 is a clone designated herein as "DNA56047-1456".

Figure 85 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:147 shown in Figure 84.

Figure 86 shows a nucleotide sequence designated herein as DNA36527 (SEQ ID NO:149).

Figure 87 shows a nucleotide sequence (SEQ ID NO:150) of a native sequence PRO1079 (UNQ536) cDNA, wherein SEQ ID NO:150 is a clone designated herein as "DNA56050-1455".

Figure 88 shows the amino acid sequence (SEQ ID NO:151) derived from the coding sequence of SEQ ID NO:150 shown in Figure 87.

Figure 89 shows a nucleotide sequence (SEQ ID NO:152) of a native sequence PRO793 (UNQ432) cDNA, wherein SEQ ID NO:152 is a clone designated herein as "DNA56110-1437".

Figure 90 shows the amino acid sequence (SEQ ID NO:153) derived from the coding sequence of SEQ ID NO:152 shown in Figure 89.

Figure 91 shows a nucleotide sequence designated herein as DNA50177 (SEQ ID NO:154).

Figure 92 shows a nucleotide sequence (SEQ ID NO:155) of a native sequence PRO1016 (UNQ499) cDNA, wherein SEQ ID NO:155 is a clone designated herein as "DNA56113-1378".

Figure 93 shows the amino acid sequence (SEQ ID NO:156) derived from the coding sequence of SEQ ID NO:155 shown in Figure 92.

Figure 94 shows a nucleotide sequence (SEQ ID NO:157) of a native sequence PRO1013 (UNQ496) cDNA, wherein SEQ ID NO:157 is a clone designated herein as "DNA56410-1414".

Figure 95 shows the amino acid sequence (SEQ ID NO:158) derived from the coding sequence of SEQ ID NO:157 shown in Figure 94.

Figure 96 shows a nucleotide sequence (SEQ ID NO:159) of a native sequence PRO937 (UNQ474) cDNA, wherein SEQ ID NO:159 is a clone designated herein as "DNA56436-1448".

Figure 97 shows the amino acid sequence (SEQ ID NO:160) derived from the coding sequence of SEQ ID NO:159 shown in Figure 96.

Figure 98 shows a nucleotide sequence (SEQ ID NO:164) of a native sequence PRO842 (UNQ473) cDNA, wherein SEQ ID NO:164 is a clone designated herein as "DNA56855-1447".

Figure 99 shows the amino acid sequence (SEQ ID NO:165) derived from the coding sequence of SEQ ID NO:164 shown in Figure 98.

Figure 100 shows a nucleotide sequence (SEQ ID NO:166) of a native sequence PRO839 (UNQ472) cDNA, wherein SEQ ID NO:166 is a clone designated herein as "DNA56859-1445".

Figure 101 shows the amino acid sequence (SEQ ID NO:167) derived from the coding sequence of SEQ ID NO:166 shown in Figure 100.

Figure 102 shows a nucleotide sequence (SEQ ID NO:168) of a native sequence PRO1180 (UNQ594) cDNA, wherein SEQ ID NO:168 is a clone designated herein as "DNA56860-1510".

Figure 103 shows the amino acid sequence (SEQ ID NO:169) derived from the coding sequence of SEQ ID NO:168 shown in Figure 102.

Figure 104 shows a nucleotide sequence (SEQ ID NO:170) of a native sequence PRO1134 (UNQ572) cDNA, wherein SEQ ID NO:170 is a clone designated herein as "DNA56865-1491".

Figure 105 shows the amino acid sequence (SEQ ID NO:171) derived from the coding sequence of SEQ ID NO:170 shown in Figure 104.

Figure 106 shows a nucleotide sequence designated herein as DNA52352 (SEQ ID NO:172).

Figure 107 shows a nucleotide sequence designated herein as DNA55725 (SEQ ID NO:173).

Figure 108 shows a nucleotide sequence (SEQ ID NO:174) of a native sequence PRO830 (UNQ470) cDNA, wherein SEQ ID NO:174 is a clone designated herein as "DNA56866-1342".

Figure 109 shows the amino acid sequence (SEQ ID NO:175) derived from the coding sequence of SEQ ID NO:174 shown in Figure 108.

Figure 110 shows a nucleotide sequence (SEQ ID NO:176) of a native sequence PRO1115 (UNQ558) cDNA, wherein SEQ ID NO:176 is a clone designated herein as "DNA56868-1478".

Figure 111 shows the amino acid sequence (SEQ ID NO:177) derived from the coding sequence of SEQ ID NO:176 shown in Figure 110.

Figure 112 shows a nucleotide sequence (SEQ ID NO:178) of a native sequence PRO1277 (UNQ647) cDNA, wherein SEQ ID NO:178 is a clone designated herein as "DNA56869-1545".

Figure 113 shows the amino acid sequence (SEQ ID NO:179) derived from the coding sequence of SEQ ID NO:178 shown in Figure 112.

Figure 114 shows a nucleotide sequence (SEQ ID NO:180) of a native sequence PRO1135 (UNQ573) cDNA, wherein SEQ ID NO:180 is a clone designated herein as "DNA56870-1492".

Figure 115 shows the amino acid sequence (SEQ ID NO:181) derived from the coding sequence of SEQ ID NO:180 shown in Figure 114.

Figure 116 shows a nucleotide sequence (SEQ ID NO:182) of a native sequence PRO1114 (UNQ557) cDNA, wherein SEQ ID NO:182 is a clone designated herein as "DNA57033-1403".

Figure 117 shows the amino acid sequence (SEQ ID NO:183) derived from the coding sequence of SEQ ID NO:182 shown in Figure 116.

Figure 118 shows a nucleotide sequence designated herein as DNA48466 (SEQ ID NO:184).

Figure 119 shows a nucleotide sequence (SEQ ID NO:188) of a native sequence PRO828 (UNQ469) cDNA, wherein SEQ ID NO:188 is a clone designated herein as "DNA57037-1444".

Figure 120 shows the amino acid sequence (SEQ ID NO:189) derived from the coding sequence of SEQ ID NO:188 shown in Figure 119.

Figure 121 shows a nucleotide sequence (SEQ ID NO:193) of a native sequence PRO1009 (UNQ493) cDNA, wherein SEQ ID NO:193 is a clone designated herein as "DNA57129-1413".

Figure 122 shows the amino acid sequence (SEQ ID NO:194) derived from the coding sequence of SEQ ID NO:193 shown in Figure 121.

Figure 123 shows a nucleotide sequence designated herein as DNA50853 (SEQ ID NO:195).

Figure 124 shows a nucleotide sequence (SEQ ID NO:196) of a native sequence PRO1007 (UNQ491) cDNA, wherein SEQ ID NO:196 is a clone designated herein as "DNA57690-1374".

Figure 125 shows the amino acid sequence (SEQ ID NO:197) derived from the coding sequence of SEQ ID NO:196 shown in Figure 124.

Figure 126 shows a nucleotide sequence (SEQ ID NO:198) of a native sequence PRO1056 (UNQ521) cDNA, wherein SEQ ID NO:198 is a clone designated herein as "DNA57693-1424".

Figure 127 shows the amino acid sequence (SEQ ID NO:199) derived from the coding sequence of SEQ ID NO:198 shown in Figure 126.

Figure 128 shows a nucleotide sequence (SEQ ID NO:200) of a native sequence PRO826 (UNQ467) cDNA, wherein SEQ ID NO:200 is a clone designated herein as "DNA57694-1341".

Figure 129 shows the amino acid sequence (SEQ ID NO:201) derived from the coding sequence of SEQ ID NO:200 shown in Figure 128.

Figure 130 shows a nucleotide sequence (SEQ ID NO:202) of a native sequence PRO819 (UNQ466) cDNA, wherein SEQ ID NO:202 is a clone designated herein as "DNA57695-1340".

Figure 131 shows the amino acid sequence (SEQ ID NO:203) derived from the coding sequence of SEQ ID NO:202 shown in Figure 130.

Figure 132 shows a nucleotide sequence (SEQ ID NO:204) of a native sequence PRO1006 (UNQ490) cDNA, wherein SEQ ID NO:204 is a clone designated herein as "DNA57699-1412".

Figure 133 shows the amino acid sequence (SEQ ID NO:205) derived from the coding sequence of SEQ ID NO:204 shown in Figure 132.

Figure 134 shows a nucleotide sequence (SEQ ID NO:206) of a native sequence PRO1112 (UNQ555) cDNA, wherein SEQ ID NO:206 is a clone designated herein as "DNA57702-1476".

Figure 135 shows the amino acid sequence (SEQ ID NO:207) derived from the coding sequence of SEQ ID NO:206 shown in Figure 134.

Figure 136 shows a nucleotide sequence (SEQ ID NO:208) of a native sequence PRO1074 (UNQ531) cDNA, wherein SEQ ID NO:208 is a clone designated herein as "DNA57704-1452".

Figure 137 shows the amino acid sequence (SEQ ID NO:209) derived from the coding sequence of SEQ ID NO:208 shown in Figure 136.

Figure 138 shows a nucleotide sequence (SEQ ID NO:210) of a native sequence PRO1005 (UNQ489) cDNA, wherein SEQ ID NO:210 is a clone designated herein as "DNA57708-1005".

Figure 139 shows the amino acid sequence (SEQ ID NO:211) derived from the coding sequence of SEQ ID NO:210 shown in Figure 138.

Figure 140 shows a nucleotide sequence (SEQ ID NO:212) of a native sequence PRO1073 (UNQ530) cDNA, wherein SEQ ID NO:212 is a clone designated herein as "DNA57710-1451".

Figure 141 shows the amino acid sequence (SEQ ID NO:213) derived from the coding sequence of SEQ ID NO:212 shown in Figure 140.

Figure 142 shows a nucleotide sequence designated herein as DNA55938 (SEQ ID NO:214).

Figure 143 shows a nucleotide sequence (SEQ ID NO:215) of a native sequence PRO1152 (UNQ582) cDNA, wherein SEQ ID NO:215 is a clone designated herein as "DNA57711-1501".

Figure 144 shows the amino acid sequence (SEQ ID NO:216) derived from the coding sequence of SEQ ID NO:215 shown in Figure 143.

Figure 145 shows a nucleotide sequence designated herein as DNA55807 (SEQ ID NO:217).

Figure 146 shows a nucleotide sequence (SEQ ID NO:218) of a native sequence PRO1136 (UNQ574) cDNA, wherein SEQ ID NO:218 is a clone designated herein as "DNA57827-1493".

Figure 147 shows the amino acid sequence (SEQ ID NO:219) derived from the coding sequence of SEQ ID NO:218 shown in Figure 146.

Figure 148 shows a nucleotide sequence (SEQ ID NO:220) of a native sequence PRO813 (UNQ465) cDNA, wherein SEQ ID NO:220 is a clone designated herein as "DNA57834-1339".

Figure 149 shows the amino acid sequence (SEQ ID NO:221) derived from the coding sequence of SEQ ID NO:220 shown in Figure 148.

Figure 150 shows a nucleotide sequence (SEQ ID NO:222) of a native sequence PRO809 (UNQ464) cDNA, wherein SEQ ID NO:222 is a clone designated herein as "DNA57836-1338".

Figure 151 shows the amino acid sequence (SEQ ID NO:223) derived from the coding sequence of SEQ ID NO:222 shown in Figure 150.

Figure 152 shows a nucleotide sequence (SEQ ID NO:224) of a native sequence PRO791 (UNQ463) cDNA, wherein SEQ ID NO:224 is a clone designated herein as "DNA57838-1337".

Figure 153 shows the amino acid sequence (SEQ ID NO:225) derived from the coding sequence of SEQ ID NO:224 shown in Figure 152.

Figure 154 shows a nucleotide sequence (SEQ ID NO:226) of a native sequence PRO1004 (UNQ488) cDNA, wherein SEQ ID NO:226 is a clone designated herein as "DNA57844-1410".

Figure 155 shows the amino acid sequence (SEQ ID NO:227) derived from the coding sequence of SEQ ID NO:226 shown in Figure 154.

Figure 156 shows a nucleotide sequence (SEQ ID NO:228) of a native sequence PRO1111 (UNQ554) cDNA, wherein SEQ ID NO:228 is a clone designated herein as "DNA58721-1475".

Figure 157 shows the amino acid sequence (SEQ ID NO:229) derived from the coding sequence of SEQ ID NO:228 shown in Figure 156.

Figure 158 shows a nucleotide sequence (SEQ ID NO:230) of a native sequence PRO1344 (UNQ699) cDNA, wherein SEQ ID NO:230 is a clone designated herein as "DNA58723-1588".

Figure 159 shows the amino acid sequence (SEQ ID NO:231) derived from the coding sequence of SEQ ID NO:230 shown in Figure 158.

Figure 160 shows a nucleotide sequence (SEQ ID NO:235) of a native sequence PRO1109 (UNQ552) cDNA, wherein SEQ ID NO:235 is a clone designated herein as "DNA58737-1473".

Figure 161 shows the amino acid sequence (SEQ ID NO:236) derived from the coding sequence of SEQ ID NO:235 shown in Figure 160.

Figure 162 shows a nucleotide sequence (SEQ ID NO:240) of a native sequence PRO1383 (UNQ719) cDNA, wherein SEQ ID NO:240 is a clone designated herein as "DNA58743-1609".

Figure 163 shows the amino acid sequence (SEQ ID NO:241) derived from the coding sequence of SEQ ID NO:240 shown in Figure 162.

Figure 164 shows a nucleotide sequence (SEQ ID NO:245) of a native sequence PRO1003 (UNQ487) cDNA, wherein SEQ ID NO:245 is a clone designated herein as "DNA58846-1409".

Figure 165 shows the amino acid sequence (SEQ ID NO:246) derived from the coding sequence of SEQ ID NO:245 shown in Figure 164.

Figure 166 shows a nucleotide sequence (SEQ ID NO:247) of a native sequence PRO1108 (UNQ551) cDNA, wherein SEQ ID NO:247 is a clone designated herein as "DNA58848-1472".

Figure 167 shows the amino acid sequence (SEQ ID NO:248) derived from the coding sequence of SEQ ID NO:247 shown in Figure 166.

Figure 168 shows a nucleotide sequence (SEQ ID NO:249) of a native sequence PRO1137 (UNQ575) cDNA, wherein SEQ ID NO:249 is a clone designated herein as "DNA58849-1494".

Figure 169 shows the amino acid sequence (SEQ ID NO:250) derived from the coding sequence of SEQ ID NO:249 shown in Figure 168.

Figure 170 shows a nucleotide sequence (SEQ ID NO:252) of a native sequence PRO1138 (UNQ576) cDNA, wherein SEQ ID NO:252 is a clone designated herein as "DNA58850-1495".

Figure 171 shows the amino acid sequence (SEQ ID NO:253) derived from the coding sequence of SEQ ID NO:252 shown in Figure 170.

Figure 172 shows a nucleotide sequence designated herein as DNA49140 (SEQ ID NO:254).

Figure 173 shows a nucleotide sequence (SEQ ID NO:255) of a native sequence PRO1054 (UNQ519) cDNA, wherein SEQ ID NO:255 is a clone designated herein as "DNA58853-1423".

Figure 174 shows the amino acid sequence (SEQ ID NO:256) derived from the coding sequence of SEQ ID NO:255 shown in Figure 173.

Figure 175 shows a nucleotide sequence (SEQ ID NO:257) of a native sequence PRO994 (UNQ518) cDNA, wherein SEQ ID NO:257 is a clone designated herein as "DNA58855-1422".

Figure 176 shows the amino acid sequence (SEQ ID NO:258) derived from the coding sequence of SEQ ID NO:257 shown in Figure 175.

Figure 177 shows a nucleotide sequence (SEQ ID NO:259) of a native sequence PRO812 (UNQ517) cDNA, wherein SEQ ID NO:259 is a clone designated herein as "DNA59205-1421".

Figure 178 shows the amino acid sequence (SEQ ID NO:260) derived from the coding sequence of SEQ ID NO:259 shown in Figure 177.

Figure 179 shows a nucleotide sequence (SEQ ID NO:261) of a native sequence PRO1069 (UNQ526) cDNA, wherein SEQ ID NO:261 is a clone designated herein as "DNA59211-1450".

Figure 180 shows the amino acid sequence (SEQ ID NO:262) derived from the coding sequence of SEQ ID NO:261 shown in Figure 179.

Figure 181 shows a nucleotide sequence (SEQ ID NO:263) of a native sequence PRO1129 (UNQ568) cDNA, wherein SEQ ID NO:263 is a clone designated herein as "DNA59213-1487".

Figure 182 shows the amino acid sequence (SEQ ID NO:264) derived from the coding sequence of SEQ ID NO:263 shown in Figure 181.

Figure 183 shows a nucleotide sequence (SEQ ID NO:265) of a native sequence PRO1068 (UNQ525) cDNA, wherein SEQ ID NO:265 is a clone designated herein as "DNA59214-1449".

Figure 184 shows the amino acid sequence (SEQ ID NO:266) derived from the coding sequence of SEQ ID NO:265 shown in Figure 183.

Figure 185 shows a nucleotide sequence (SEQ ID NO:267) of a native sequence PRO1066 (UNQ524) cDNA, wherein SEQ ID NO:267 is a clone designated herein as "DNA59215-1425".

Figure 186 shows the amino acid sequence (SEQ ID NO:268) derived from the coding sequence of SEQ ID NO:267 shown in Figure 185.

Figure 187 shows a nucleotide sequence (SEQ ID NO:269) of a native sequence PRO1184 (UNQ598) cDNA, wherein SEQ ID NO:269 is a clone designated herein as "DNA59220-1514".

Figure 188 shows the amino acid sequence (SEQ ID NO:270) derived from the coding sequence of SEQ ID NO:269 shown in Figure 187.

Figure 189 shows a nucleotide sequence (SEQ ID NO:271) of a native sequence PRO1360 (UNQ709) cDNA, wherein SEQ ID NO:271 is a clone designated herein as "DNA59488-1603".

Figure 190 shows the amino acid sequence (SEQ ID NO:272) derived from the coding sequence of SEQ ID NO:271 shown in Figure 189.

Figure 191 shows a nucleotide sequence (SEQ ID NO:273) of a native sequence PRO1029 (UNQ514) cDNA, wherein SEQ ID NO:273 is a clone designated herein as "DNA59493-1420".

Figure 192 shows the amino acid sequence (SEQ ID NO:274) derived from the coding sequence of SEQ ID NO:273 shown in Figure 191.

Figure 193 shows a nucleotide sequence (SEQ ID NO:275) of a native sequence PRO1139 (UNQ577) cDNA, wherein SEQ ID NO:275 is a clone designated herein as "DNA59497-1496".

5 Figure 194 shows the amino acid sequence (SEQ ID NO:276) derived from the coding sequence of SEQ ID NO:275 shown in Figure 193.

Figure 195 shows a nucleotide sequence (SEQ ID NO:277) of a native sequence PRO1309 (UNQ675) cDNA, wherein SEQ ID NO:277 is a clone designated herein as "DNA59588-1571".

Figure 196 shows the amino acid sequence (SEQ ID NO:278) derived from the coding sequence of SEQ ID NO:277 shown in Figure 195.

10 Figure 197 shows a nucleotide sequence (SEQ ID NO:280) of a native sequence PRO1028 (UNQ513) cDNA, wherein SEQ ID NO:280 is a clone designated herein as "DNA59603-1419".

Figure 198 shows the amino acid sequence (SEQ ID NO:281) derived from the coding sequence of SEQ ID NO:280 shown in Figure 197.

15 Figure 199 shows a nucleotide sequence (SEQ ID NO:282) of a native sequence PRO1027 (UNQ512) cDNA, wherein SEQ ID NO:282 is a clone designated herein as "DNA59605-1418".

Figure 200 shows the amino acid sequence (SEQ ID NO:283) derived from the coding sequence of SEQ ID NO:282 shown in Figure 199.

Figure 201 shows a nucleotide sequence (SEQ ID NO:284) of a native sequence PRO1107 (UNQ550) cDNA, wherein SEQ ID NO:284 is a clone designated herein as "DNA59606-1471".

20 Figure 202 shows the amino acid sequence (SEQ ID NO:285) derived from the coding sequence of SEQ ID NO:284 shown in Figure 201.

Figure 203 shows a nucleotide sequence (SEQ ID NO:286) of a native sequence PRO1140 (UNQ578) cDNA, wherein SEQ ID NO:286 is a clone designated herein as "DNA59607-1497".

25 Figure 204 shows the amino acid sequence (SEQ ID NO:287) derived from the coding sequence of SEQ ID NO:286 shown in Figure 203.

Figure 205 shows a nucleotide sequence (SEQ ID NO:288) of a native sequence PRO1106 (UNQ549) cDNA, wherein SEQ ID NO:288 is a clone designated herein as "DNA59609-1470".

Figure 206 shows the amino acid sequence (SEQ ID NO:289) derived from the coding sequence of SEQ ID NO:288 shown in Figure 205.

30 Figure 207 shows a nucleotide sequence (SEQ ID NO:290) of a native sequence PRO1291 (UNQ659) cDNA, wherein SEQ ID NO:290 is a clone designated herein as "DNA59610-1556".

Figure 208 shows the amino acid sequence (SEQ ID NO:291) derived from the coding sequence of SEQ ID NO:290 shown in Figure 207.

35 Figure 209 shows a nucleotide sequence (SEQ ID NO:292) of a native sequence PRO1105 (UNQ548) cDNA, wherein SEQ ID NO:292 is a clone designated herein as "DNA59612-1466".

Figure 210 shows the amino acid sequence (SEQ ID NO:293) derived from the coding sequence of SEQ ID NO:292 shown in Figure 209.

Figure 211 shows a nucleotide sequence (SEQ ID NO:294) of a native sequence PRO511 (UNQ511) cDNA, wherein SEQ ID NO:294 is a clone designated herein as "DNA59613-1417".

Figure 212 shows the amino acid sequence (SEQ ID NO:295) derived from the coding sequence of SEQ ID NO:294 shown in Figure 211.

5 Figure 213 shows a nucleotide sequence (SEQ ID NO:296) of a native sequence PRO1104 (UNQ547) cDNA, wherein SEQ ID NO:296 is a clone designated herein as "DNA59616-1465".

Figure 214 shows the amino acid sequence (SEQ ID NO:297) derived from the coding sequence of SEQ ID NO:296 shown in Figure 213.

Figure 215 shows a nucleotide sequence (SEQ ID NO:298) of a native sequence PRO1100 (UNQ546) cDNA, wherein SEQ ID NO:298 is a clone designated herein as "DNA59619-1464".

10 Figure 216 shows the amino acid sequence (SEQ ID NO:299) derived from the coding sequence of SEQ ID NO:298 shown in Figure 215.

Figure 217 shows a nucleotide sequence (SEQ ID NO:300) of a native sequence PRO836 (UNQ545) cDNA, wherein SEQ ID NO:300 is a clone designated herein as "DNA59620-1463".

15 Figure 218 shows the amino acid sequence (SEQ ID NO:301) derived from the coding sequence of SEQ ID NO:300 shown in Figure 217.

Figure 219 shows a nucleotide sequence (SEQ ID NO:302) of a native sequence PRO1141 (UNQ579) cDNA, wherein SEQ ID NO:302 is a clone designated herein as "DNA59625-1498".

Figure 220 shows the amino acid sequence (SEQ ID NO:303) derived from the coding sequence of SEQ ID NO:302 shown in Figure 219.

20 Figure 221 shows a nucleotide sequence designated herein as DNA33128 (SEQ ID NO:304).

Figure 222 shows a nucleotide sequence designated herein as DNA34256 (SEQ ID NO:305).

Figure 223 shows a nucleotide sequence designated herein as DNA47941 (SEQ ID NO:306).

Figure 224 shows a nucleotide sequence designated herein as DNA54389 (SEQ ID NO:307).

25 Figure 225 shows a nucleotide sequence (SEQ ID NO:308) of a native sequence PRO1132 (UNQ570) cDNA, wherein SEQ ID NO:308 is a clone designated herein as "DNA59767-1489".

Figure 226 shows the amino acid sequence (SEQ ID NO:309) derived from the coding sequence of SEQ ID NO:308 shown in Figure 225.

Figure 227 shows a nucleotide sequence (SEQ ID NO:313) of a native sequence PRO1346 cDNA, wherein SEQ ID NO:313 is a clone designated herein as "DNA59776-1600".

30 Figure 228 shows the amino acid sequence (SEQ ID NO:314) derived from the coding sequence of SEQ ID NO:313 shown in Figure 227.

Figure 229 shows a nucleotide sequence (SEQ ID NO:318) of a native sequence PRO1131 (UNQ569) cDNA, wherein SEQ ID NO:318 is a clone designated herein as "DNA59777-1480".

35 Figure 230 shows the amino acid sequence (SEQ ID NO:319) derived from the coding sequence of SEQ ID NO:318 shown in Figure 229.

Figure 231 shows a nucleotide sequence designated herein as DNA43546 (SEQ ID NO:320).

Figure 232 shows a nucleotide sequence (SEQ ID NO:325) of a native sequence PRO1281 (UNQ651) cDNA, wherein SEQ ID NO:325 is a clone designated herein as "DNA59820-1549".

Figure 233 shows the amino acid sequence (SEQ ID NO:326) derived from the coding sequence of SEQ ID NO:325 shown in Figure 232.

Figure 234 shows a nucleotide sequence (SEQ ID NO:333) of a native sequence PRO1064 (UNQ111) cDNA, wherein SEQ ID NO:333 is a clone designated herein as "DNA59827-1426".

Figure 235 shows the amino acid sequence (SEQ ID NO:334) derived from the coding sequence of SEQ ID NO:333 shown in Figure 234.

Figure 236 shows a nucleotide sequence designated herein as DNA45288 (SEQ ID NO:335).

Figure 237 shows a nucleotide sequence (SEQ ID NO:339) of a native sequence PRO1379 (UNQ716) cDNA, wherein SEQ ID NO:339 is a clone designated herein as "DNA59828-1608".

Figure 238 shows the amino acid sequence (SEQ ID NO:340) derived from the coding sequence of SEQ ID NO:339 shown in Figure 237.

Figure 239 shows a nucleotide sequence (SEQ ID NO:344) of a native sequence PRO844 (UNQ544) cDNA, wherein SEQ ID NO:344 is a clone designated herein as "DNA59838-1462".

Figure 240 shows the amino acid sequence (SEQ ID NO:345) derived from the coding sequence of SEQ ID NO:344 shown in Figure 239.

Figure 241 shows a nucleotide sequence (SEQ ID NO:346) of a native sequence PRO848 (UNQ543) cDNA, wherein SEQ ID NO:346 is a clone designated herein as "DNA59839-1461".

Figure 242 shows the amino acid sequence (SEQ ID NO:347) derived from the coding sequence of SEQ ID NO:346 shown in Figure 241.

Figure 243 shows a nucleotide sequence (SEQ ID NO:348) of a native sequence PRO1097 (UNQ542) cDNA, wherein SEQ ID NO:348 is a clone designated herein as "DNA59841-1460".

Figure 244 shows the amino acid sequence (SEQ ID NO:349) derived from the coding sequence of SEQ ID NO:348 shown in Figure 243.

Figure 245 shows a nucleotide sequence (SEQ ID NO:350) of a native sequence PRO1153 (UNQ583) cDNA, wherein SEQ ID NO:350 is a clone designated herein as "DNA59842-1502".

Figure 246 shows the amino acid sequence (SEQ ID NO:351) derived from the coding sequence of SEQ ID NO:350 shown in Figure 245.

Figure 247 shows a nucleotide sequence (SEQ ID NO:352) of a native sequence PRO1154 (UNQ584) cDNA, wherein SEQ ID NO:352 is a clone designated herein as "DNA59846-1503".

Figure 248 shows the amino acid sequence (SEQ ID NO:353) derived from the coding sequence of SEQ ID NO:352 shown in Figure 247.

Figure 249 shows a nucleotide sequence (SEQ ID NO:354) of a native sequence PRO1181 (UNQ595) cDNA, wherein SEQ ID NO:354 is a clone designated herein as "DNA59847-1511".

Figure 250 shows the amino acid sequence (SEQ ID NO:355) derived from the coding sequence of SEQ ID NO:354 shown in Figure 249.

Figure 251 shows a nucleotide sequence (SEQ ID NO:356) of a native sequence PRO1182 (UNQ596) cDNA, wherein SEQ ID NO:356 is a clone designated herein as "DNA59848-1512".

Figure 252 shows the amino acid sequence (SEQ ID NO:357) derived from the coding sequence of SEQ ID NO:356 shown in Figure 251.

5 Figure 253 shows a nucleotide sequence (SEQ ID NO:358) of a native sequence PRO1155 (UNQ585) cDNA, wherein SEQ ID NO:358 is a clone designated herein as "DNA59849-1504".

Figure 254 shows the amino acid sequence (SEQ ID NO:359) derived from the coding sequence of SEQ ID NO:358 shown in Figure 253.

Figure 255 shows a nucleotide sequence (SEQ ID NO:360) of a native sequence PRO1156 (UNQ586) cDNA, wherein SEQ ID NO:360 is a clone designated herein as "DNA59853-1505".

10 Figure 256 shows the amino acid sequence (SEQ ID NO:361) derived from the coding sequence of SEQ ID NO:360 shown in Figure 255.

Figure 257 shows a nucleotide sequence (SEQ ID NO:362) of a native sequence PRO1098 (UNQ541) cDNA, wherein SEQ ID NO:362 is a clone designated herein as "DNA59854-1459".

15 Figure 258 shows the amino acid sequence (SEQ ID NO:363) derived from the coding sequence of SEQ ID NO:362 shown in Figure 257.

Figure 259 shows a nucleotide sequence (SEQ ID NO:364) of a native sequence PRO1127 (UNQ565) cDNA, wherein SEQ ID NO:364 is a clone designated herein as "DNA60283-1484".

Figure 260 shows the amino acid sequence (SEQ ID NO:365) derived from the coding sequence of SEQ ID NO:364 shown in Figure 259.

20 Figure 261 shows a nucleotide sequence (SEQ ID NO:366) of a native sequence PRO1126 (UNQ564) cDNA, wherein SEQ ID NO:366 is a clone designated herein as "DNA60615-1483".

Figure 262 shows the amino acid sequence (SEQ ID NO:367) derived from the coding sequence of SEQ ID NO:366 shown in Figure 261.

25 Figure 263 shows a nucleotide sequence (SEQ ID NO:368) of a native sequence PRO1125 (UNQ563) cDNA, wherein SEQ ID NO:368 is a clone designated herein as "DNA60619-1482".

Figure 264 shows the amino acid sequence (SEQ ID NO:369) derived from the coding sequence of SEQ ID NO:368 shown in Figure 263.

Figure 265 shows a nucleotide sequence (SEQ ID NO:370) of a native sequence PRO1186 (UNQ600) cDNA, wherein SEQ ID NO:370 is a clone designated herein as "DNA60621-1516".

30 Figure 266 shows the amino acid sequence (SEQ ID NO:371) derived from the coding sequence of SEQ ID NO:370 shown in Figure 265.

Figure 267 shows a nucleotide sequence (SEQ ID NO:372) of a native sequence PRO1198 (UNQ611) cDNA, wherein SEQ ID NO:372 is a clone designated herein as "DNA60622-1525".

35 Figure 268 shows the amino acid sequence (SEQ ID NO:373) derived from the coding sequence of SEQ ID NO:372 shown in Figure 267.

Figure 269 shows a nucleotide sequence (SEQ ID NO:374) of a native sequence PRO1158 (UNQ588) cDNA, wherein SEQ ID NO:374 is a clone designated herein as "DNA60625-1507".

Figure 270 shows the amino acid sequence (SEQ ID NO:375) derived from the coding sequence of SEQ ID NO:374 shown in Figure 269.

Figure 271 shows a nucleotide sequence (SEQ ID NO:376) of a native sequence PRO1159 (UNQ589) cDNA, wherein SEQ ID NO:376 is a clone designated herein as "DNA60627-1508".

5 Figure 272 shows the amino acid sequence (SEQ ID NO:377) derived from the coding sequence of SEQ ID NO:376 shown in Figure 271.

Figure 273 shows a nucleotide sequence (SEQ ID NO:378) of a native sequence PRO1124 (UNQ562) cDNA, wherein SEQ ID NO:378 is a clone designated herein as "DNA60629-1481".

Figure 274 shows the amino acid sequence (SEQ ID NO:379) derived from the coding sequence of SEQ ID NO:378 shown in Figure 273.

10 Figure 275 shows a nucleotide sequence (SEQ ID NO:380) of a native sequence PRO1287 (UNQ656) cDNA, wherein SEQ ID NO:380 is a clone designated herein as "DNA61755-1554".

Figure 276 shows the amino acid sequence (SEQ ID NO:381) derived from the coding sequence of SEQ ID NO:380 shown in Figure 275.

15 Figure 277 shows a nucleotide sequence (SEQ ID NO:386) of a native sequence PRO1312 (UNQ678) cDNA, wherein SEQ ID NO:386 is a clone designated herein as "DNA61873-1574".

Figure 278 shows the amino acid sequence (SEQ ID NO:387) derived from the coding sequence of SEQ ID NO:386 shown in Figure 277.

Figure 279 shows a nucleotide sequence (SEQ ID NO:388) of a native sequence PRO1192 (UNQ606) cDNA, wherein SEQ ID NO:388 is a clone designated herein as "DNA62814-1521".

20 Figure 280 shows the amino acid sequence (SEQ ID NO:389) derived from the coding sequence of SEQ ID NO:388 shown in Figure 279.

Figure 281 shows a nucleotide sequence (SEQ ID NO:393) of a native sequence PRO1160 (UNQ590) cDNA, wherein SEQ ID NO:393 is a clone designated herein as "DNA62872-1509".

25 Figure 282 shows the amino acid sequence (SEQ ID NO:394) derived from the coding sequence of SEQ ID NO:393 shown in Figure 281.

Figure 283 shows a nucleotide sequence (SEQ ID NO:398) of a native sequence PRO1187 (UNQ601) cDNA, wherein SEQ ID NO:398 is a clone designated herein as "DNA62876-1517".

Figure 284 shows the amino acid sequence (SEQ ID NO:399) derived from the coding sequence of SEQ ID NO:398 shown in Figure 283.

30 Figure 285 shows a nucleotide sequence (SEQ ID NO:400) of a native sequence PRO1185 (UNQ599) cDNA, wherein SEQ ID NO:400 is a clone designated herein as "DNA62881-1515".

Figure 286 shows the amino acid sequence (SEQ ID NO:401) derived from the coding sequence of SEQ ID NO:400 shown in Figure 285.

35 Figure 287 shows a nucleotide sequence (SEQ ID NO:402) of a native sequence PRO1345 (UNQ700) cDNA, wherein SEQ ID NO:402 is a clone designated herein as "DNA64852-1589".

Figure 288 shows the amino acid sequence (SEQ ID NO:403) derived from the coding sequence of SEQ ID NO:402 shown in Figure 287.

Figure 289 shows a nucleotide sequence (SEQ ID NO:407) of a native sequence PRO1245 (UNQ629) cDNA, wherein SEQ ID NO:407 is a clone designated herein as "DNA64884-1527".

Figure 290 shows the amino acid sequence (SEQ ID NO:408) derived from the coding sequence of SEQ ID NO:407 shown in Figure 289.

Figure 291 shows a nucleotide sequence (SEQ ID NO:409) of a native sequence PRO1358 (UNQ707) cDNA, wherein SEQ ID NO:409 is a clone designated herein as "DNA64890-1612".

Figure 292 shows the amino acid sequence (SEQ ID NO:410) derived from the coding sequence of SEQ ID NO:409 shown in Figure 291.

Figure 293 shows a nucleotide sequence (SEQ ID NO:411) of a native sequence PRO1195 (UNQ608) cDNA, wherein SEQ ID NO:411 is a clone designated herein as "DNA65412-1523".

Figure 294 shows the amino acid sequence (SEQ ID NO:412) derived from the coding sequence of SEQ ID NO:411 shown in Figure 293.

Figure 295 shows a nucleotide sequence (SEQ ID NO:413) of a native sequence PRO1270 (UNQ640) cDNA, wherein SEQ ID NO:413 is a clone designated herein as "DNA66308-1537".

Figure 296 shows the amino acid sequence (SEQ ID NO:414) derived from the coding sequence of SEQ ID NO:413 shown in Figure 295.

Figure 297 shows a nucleotide sequence (SEQ ID NO:415) of a native sequence PRO1271 (UNQ641) cDNA, wherein SEQ ID NO:415 is a clone designated herein as "DNA66309-1538".

Figure 298 shows the amino acid sequence (SEQ ID NO:416) derived from the coding sequence of SEQ ID NO:415 shown in Figure 297.

Figure 299 shows a nucleotide sequence (SEQ ID NO:417) of a native sequence PRO1375 (UNQ712) cDNA, wherein SEQ ID NO:417 is a clone designated herein as "DNA67004-1614".

Figure 300 shows the amino acid sequence (SEQ ID NO:418) derived from the coding sequence of SEQ ID NO:417 shown in Figure 299.

Figure 301 shows a nucleotide sequence (SEQ ID NO:419) of a native sequence PRO1385 (UNQ720) cDNA, wherein SEQ ID NO:419 is a clone designated herein as "DNA68869-1610".

Figure 302 shows the amino acid sequence (SEQ ID NO:420) derived from the coding sequence of SEQ ID NO:419 shown in Figure 301.

Figure 303 shows a nucleotide sequence (SEQ ID NO:421) of a native sequence PRO1387 (UNQ722) cDNA, wherein SEQ ID NO:421 is a clone designated herein as "DNA68872-1620".

Figure 304 shows the amino acid sequence (SEQ ID NO:422) derived from the coding sequence of SEQ ID NO:421 shown in Figure 303.

Figure 305 shows a nucleotide sequence (SEQ ID NO:423) of a native sequence PRO1384 (UNQ721) cDNA, wherein SEQ ID NO:423 is a clone designated herein as "DNA71159-1617".

Figure 306 shows the amino acid sequence (SEQ ID NO:424) derived from the coding sequence of SEQ ID NO:423 shown in Figure 305.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTSI. Definitions

The terms "PRO polypeptide" and "PRO" as used herein and when immediately followed by a numerical designation refer to various polypeptides, wherein the complete designation (i.e., PRO/number) refers to specific polypeptide sequences as described herein. The terms "PRO/number polypeptide" and "PRO/number" wherein the term "number" is provided an actual numerical designation as used herein encompass native sequence polypeptides and polypeptide variants (which are further defined herein). The PRO polypeptides described herein may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods.

A "native sequence PRO polypeptide" comprises a polypeptide having the same amino acid sequence as the corresponding PRO polypeptide derived from nature. Such native sequence PRO polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific PRO polypeptide (e.g., an extracellular domain sequence), naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In various embodiments of the invention, the native sequence PRO281 polypeptide is a mature or full-length native sequence PRO281 polypeptide comprising amino acids 1 to 345 of Figure 2 (SEQ ID NO:2), the native sequence PRO276 is a full-length or mature native sequence PRO276 comprising amino acids 1 through 251 of Figure 4 (SEQ ID NO:6), the native sequence PRO189 is a full-length or mature native sequence PRO189 comprising amino acids 1 through 367 of Figure 6 (SEQ ID NO:8), the native sequence PRO190 polypeptide is a full-length or mature native sequence PRO190 polypeptide comprising amino acids 1 through 424 of Figure 9 (SEQ ID NO:14), the native sequence PRO341 is a mature or full-length native sequence PRO341 comprising amino acids 1 to 458 of Figure 12 (SEQ ID NO:20), the native sequence PRO180 is a full-length or mature native sequence PRO180 comprising amino acids 1 through 266 of Figure 15 (SEQ ID NO:23), the native sequence PRO194 polypeptide is a mature or full-length native sequence PRO194 polypeptide comprising amino acids 1 to 264 of Figure 18 (SEQ ID NO:28), the native sequence PRO203 polypeptide is a mature or full-length native sequence PRO203 polypeptide comprising amino acids 1 to 347 of Figure 20 (SEQ ID NO:30), the native sequence PRO290 is a full-length or mature native sequence PRO290 comprising amino acids 1 through 1003 of Figure 23 (SEQ ID NO:33), the native sequence PRO874 polypeptide comprises amino acids 1 to 321 of Figure 25 (SEQ ID NO:36), the native sequence PRO710 polypeptide is a mature or full-length native sequence PRO710 polypeptide comprising amino acids 1 to 566 of Figure 27 (SEQ ID NO:41), the native sequence PRO1151 is a mature or full-length native sequence PRO1151 comprising amino acids 1 to 259 of Figure 30 (SEQ ID NO:47), the native sequence PRO1282 is a full-length or mature native sequence PRO1282 comprising amino acids 1 or about 24 through 673 of Figure 32 (SEQ ID NO:52), the native sequence PRO358 is a mature or full-length native sequence PRO358 polypeptide comprising amino acids 1 to 811 of Figure 34 (SEQ ID NO:57), the native sequence PRO1310 is a full-length or mature native sequence PRO1310 comprising amino acids 1 through 765 of Figure 36 (SEQ ID NO:62), the native sequence PRO698 polypeptide is a mature or full-length native sequence PRO698 polypeptide comprising amino acids 1 to 510 of Figure 38 (SEQ ID NO:67), the native

sequence PRO732 polypeptide is a mature or full-length native sequence PRO732 polypeptide comprising amino acids 1 to 453 of Figure 41 (SEQ ID NO:73), the native sequence PRO1120 is a full-length or mature native sequence PRO1120 comprising amino acids 1 or about 18 to 867 of Figure 47 (SEQ ID NO:84), the native sequence PRO537 is a mature or full-length native sequence PRO537 comprising amino acids 1 to 115 of Figure 49 (SEQ ID NO:95), the native sequence PRO536 is a mature or full-length native sequence PRO536 comprising amino acids 1 to 313 of Figure 51 (SEQ ID NO:97), the native sequence PRO535 is a mature or full-length native sequence PRO535 comprising amino acids 1 to 201 of Figure 53 (SEQ ID NO:99), the native sequence PRO718 polypeptide is a mature or full-length native sequence PRO718 polypeptide comprising amino acids 1 to 157 of Figure 57 (SEQ ID NO:103), the native sequence PRO872 polypeptide is a mature or full-length native sequence PRO872 polypeptide comprising amino acids 1 to 610 of Figure 64 (SEQ ID NO:113), the native sequence PRO1063 polypeptide is a mature or full-length native sequence PRO1063 polypeptide comprising amino acids 1 to 301 of Figure 66 (SEQ ID NO:115), the native sequence PRO619 is a full-length or mature native sequence PRO619 comprising amino acids 1 or about 21 through 123 of Figure 68 (SEQ ID NO:117), the native sequence PRO943 is a mature or full-length native sequence PRO943 comprising amino acids 1 to 504 of Figure 70 (SEQ ID NO:119), the native sequence PRO1188 is a full-length or mature native sequence PRO1188 comprising amino acids 1 or about 22 to 1184 of Figure 72 (SEQ ID NO:124), the native sequence PRO1133 is a full-length or mature native sequence PRO1133 comprising amino acids 1 or about 19 through 438 of Figure 74 (SEQ ID NO:129), the native sequence PRO784 is a mature or full-length native sequence PRO784 comprising amino acids 16 to 228 of Figure 76 (SEQ ID NO:135), the native sequence PRO783 polypeptide is a mature or full-length native sequence PRO783 polypeptide comprising amino acids 1 to 489 of Figure 79 (SEQ ID NO:138), the native sequence PRO820 is a full-length or mature native sequence PRO820 comprising amino acids 1 or 16 through 124 of Figure 83 (SEQ ID NO:146), the native sequence PRO1080 is a full-length or mature native sequence PRO1080 comprising amino acids 1 or 23 through 358 of Figure 85 (SEQ ID NO:148), the native sequence PRO1079 is a full-length or mature native sequence PRO1079 comprising amino acids 1 or about 30 to 226 of Figure 88 (SEQ ID NO:151), the native sequence PRO793 is a mature or full-length native sequence PRO793 comprising amino acids 1 to 138 of Figure 90 (SEQ ID NO:153), the native sequence PRO1016 is a full-length or mature native sequence PRO1016 comprising amino acids 1 or 19 through 378 of Figure 93 (SEQ ID NO:156), the native sequence PRO1013 polypeptide is a full-length or mature native sequence PRO1013 polypeptide comprising amino acids 1 or 20 through 409 of Figure 95 (SEQ ID NO:158), the native sequence PRO937 polypeptide is a mature or full-length native sequence PRO937 polypeptide comprising amino acids 1 to 556 of Figure 97 (SEQ ID NO:160), the native sequence PRO842 is a full-length or mature native sequence PRO842 comprising amino acids 1 or about 23 to 119 of Figure 99 (SEQ ID NO:165), the native sequence PRO839 is a full-length or mature native sequence PRO839 comprising amino acids 1 or about 24 to 87 of Figure 101 (SEQ ID NO:167), the native sequence PRO1180 polypeptide is a mature or full-length native sequence PRO1180 polypeptide comprising amino acids 1 to 277 of Figure 103 (SEQ ID NO:169), the native sequence PRO1134 is a mature or full-length native sequence PRO1134 comprising amino acids 1 to 371 of Figure 105 (SEQ ID NO:171), the native sequence PRO830 is a mature or full-length native sequence PRO830 comprising amino acids 1 to 87 of Figure 109 (SEQ ID NO:175), the native sequence

PRO1115 is a full-length or mature native sequence PRO1115 comprising amino acids 1 or about 21 to 445 of Figure 111 (SEQ ID NO:177), the native sequence PRO1277 is a full-length or mature native sequence PRO1277 comprising amino acids 1 or about 27 to 678 of Figure 113 (SEQ ID NO:179), the native sequence PRO1135 polypeptide is a mature or full-length native sequence PRO1135 polypeptide comprising amino acids 1 to 541 of Figure 115 (SEQ ID NO:181), the native sequence PRO1114 interferon receptor is a mature or full-length native sequence PRO1114 interferon receptor comprising amino acids 1 to 311 of Figure 118 (SEQ ID NO:184), the native sequence PRO828 polypeptide is a mature or full-length native sequence PRO828 polypeptide comprising amino acids 1 to 187 of Figure 120 (SEQ ID NO:189), the native sequence PRO1009 is a full-length or mature native sequence PRO1009 comprising amino acids 1 or 23 to 615 of Figure 122 (SEQ ID NO:194), the native sequence PRO1007 polypeptide is a full-length or mature native sequence PRO1007 polypeptide comprising amino acids 1 or 31 through 346 of Figure 125 (SEQ ID NO:197), the native sequence PRO1056 is a mature or full-length native sequence PRO1056 comprising amino acids 1 to 120 of Figure 127 (SEQ ID NO:199), the native sequence PRO826 is a mature or full-length native sequence PRO826 comprising amino acids 1 to 99 of Figure 129 (SEQ ID NO:201), the native sequence PRO819 is a mature or full-length native sequence PRO819 comprising amino acids 1 to 52 of Figure 131 (SEQ ID NO:203), the native sequence PRO1006 is a full-length or mature native sequence PRO1006 comprising amino acids 1 or 24 through 392 of Figure 133 (SEQ ID NO:205), the native sequence PRO1112 polypeptide is a full-length or mature native sequence PRO1112 polypeptide comprising amino acids 1 or 14 through 262 of Figure 135 (SEQ ID NO:207), the native sequence PRO1074 polypeptide is a mature or full-length native sequence PRO1074 polypeptide comprising amino acids 1 to 331 of Figure 137 (SEQ ID NO:209), the native sequence PRO1005 is a full-length or mature native sequence PRO1005 comprising amino acids 1 or about 21 to 185 of Figure 139 (SEQ ID NO:211), the native sequence PRO1073 is a full-length or mature native sequence PRO1073 comprising amino acids 1 or about 32 to 299 of Figure 141 (SEQ ID NO:213), the native sequence PRO1152 is a mature or full-length native sequence PRO1152 comprising amino acids 1 to 479 of Figure 144 (SEQ ID NO:216), the native sequence PRO1136 is a mature or full-length native sequence PRO1136 comprising amino acids 1 to 632 of Figure 147 (SEQ ID NO:219), the native sequence PRO813 polypeptide is a mature or full-length native sequence PRO813 polypeptide comprising amino acids 1 to 76 of Figure 149 (SEQ ID NO:221), the native sequence PRO809 is a full-length or mature native sequence PRO809 comprising amino acids 1 or 19 through 265 of Figure 151 (SEQ ID NO:223), the native sequence PRO791 is a full-length or mature native sequence PRO791 comprising amino acids 1 or 26 through 246 of Figure 153 (SEQ ID NO:225), the native sequence PRO1004 is a full-length or mature native sequence PRO1004 comprising amino acids 1 or about 25 through 115 of Figure 155 (SEQ ID NO:227), the native sequence PRO1111 is a full-length or mature native sequence PRO1111 comprising amino acids 1 through 653 of Figure 157 (SEQ ID NO:229), the native sequence PRO1344 is a mature or full-length native sequence PRO1344 comprising amino acids 1 to 720 of Figure 159 (SEQ ID NO:231), the native sequence PRO1109 is a mature or full-length native sequence PRO1109 comprising amino acids 1 to 344 of Figure 161 (SEQ ID NO:236), the native sequence PRO1383 is a mature or full-length native sequence PRO1383 comprising amino acids 1 to 423 of Figure 163 (SEQ ID NO:241), the native sequence PRO1003 polypeptide is a mature or full-length native sequence PRO1003 polypeptide comprising amino acids

1 to 84 of Figure 165 (SEQ ID NO:246), the native sequence PRO1108 polypeptide is a mature or full-length native sequence PRO1108 polypeptide comprising amino acids 1 to 456 of Figure 167 (SEQ ID NO:248), the native sequence PRO1137 polypeptide is a mature or full-length native sequence PRO1137 polypeptide comprising amino acids 1 to 240 of Figure 169 (SEQ ID NO:250), the native sequence PRO1138 polypeptide is a mature or full-length native sequence PRO1138 polypeptide comprising amino acids 1 to 335 of Figure 171 (SEQ ID NO:253), the native sequence PRO1054 is a mature or full-length native sequence PRO1054 comprising amino acids 1 to 180 of Figure 174 (SEQ ID NO:256), the native sequence PRO994 is a mature or full-length native sequence PRO994 comprising amino acids 1 to 229 of Figure 176 (SEQ ID NO:258), the native sequence PRO812 is a mature or full-length native sequence PRO812 comprising amino acids 1 to 83 of Figure 178 (SEQ ID NO:260), the native sequence PRO1069 polypeptide is a mature or full-length native sequence PRO1069 polypeptide comprising amino acids 1 to 89 of Figure 180 (SEQ ID NO:262), the native sequence PRO1129 polypeptide is a mature or full-length native sequence PRO1129 polypeptide comprising amino acids 1 to 524 of Figure 182 (SEQ ID NO:264), the native sequence PRO1068 is a full-length or mature native sequence PRO1068 comprising amino acids 1 or about 21 to 124 of Figure 184 (SEQ ID NO:266), the native sequence PRO1066 polypeptide is a mature or full-length native sequence PRO1066 polypeptide comprising amino acids 1 to 117 of Figure 186 (SEQ ID NO:268), the native sequence PRO1184 polypeptide is a full-length or mature native sequence PRO1184 polypeptide comprising amino acids 1 or 39 through 142 of Figure 188 (SEQ ID NO:270), the native sequence PRO1360 is a full-length or mature native sequence PRO1360 comprising amino acids 1 or about 30 through 285 of Figure 190 (SEQ ID NO:272), the native sequence PRO1029 is a mature or full-length native sequence PRO1029 comprising amino acids 1 to 86 of Figure 192 (SEQ ID NO:274), the native sequence PRO1139 is a mature or full-length native sequence PRO1139 polypeptide comprising amino acids 1 to 131 or 29-131 of Figure 194 (SEQ ID NO:276), the native sequence PRO1309 is a full-length or mature native sequence PRO1309 comprising amino acids 1 or about 35 through 522 of Figure 196 (SEQ ID NO:278), the native sequence PRO1028 polypeptide is a full-length or mature native sequence PRO1028 polypeptide comprising amino acids 1 or 20 through 197 of Figure 198 (SEQ ID NO:281), the native sequence PRO1027 is a full-length or mature native sequence PRO1027 comprising amino acids 1 or 34 through 77 of Figure 200 (SEQ ID NO:283), the native sequence PRO1107 polypeptide is a full-length or mature native sequence PRO1107 polypeptide comprising amino acids 1 or 23 through 477 of Figure 202 (SEQ ID NO:285), the native sequence PRO1140 polypeptide is a mature or full-length native sequence PRO1140 polypeptide comprising amino acids 1 to 255 of Figure 204 (SEQ ID NO:287), the native sequence PRO1106 polypeptide is a full-length or mature native sequence PRO1106 polypeptide comprising amino acids 1 or 17 through 469 of Figure 206 (SEQ ID NO:289), the native sequence PRO1291 is a mature or full-length native sequence PRO1291 comprising amino acids 1 to 282 of Figure 208 (SEQ ID NO:291), the native sequence PRO1105 polypeptide is a full-length or mature native sequence PRO1105 polypeptide comprising amino acids 1 or 20 through 180 of Figure 210 (SEQ ID NO:293), the native sequence PRO1026 is a full-length or mature native sequence PRO1026 comprising amino acids 1 or 26 through 237 of Figure 212 (SEQ ID NO:295), the native sequence PRO1104 is a full-length or mature native sequence PRO1104 comprising amino acids 1 or about 23 through 341 of Figure 214 (SEQ ID NO:297), the native sequence PRO1100 is a full-length or mature native

sequence PRO1100 comprising amino acids 1 or 21 through 320 of Figure 216 (SEQ ID NO:299), the native sequence PRO836 is a full-length or mature native sequence PRO836 comprising amino acids 1 or 30 through 461 of Figure 218 (SEQ ID NO:301), the native sequence PRO1141 is a mature or full-length native sequence PRO1141 comprising amino acids 1 to 247 of Figure 220 (SEQ ID NO:303), the native sequence PRO1132 is a full-length or mature native sequence PRO1132 comprising amino acids 1 or about 23 through 293 of Figure 226 (SEQ ID NO:309), the native sequence NL7 is a mature or full-length native sequence NL7 comprising amino acids from about position 51 to about position 461 of Figure 228 (SEQ ID NO:314), the native sequence PRO1131 is a full-length or mature native sequence PRO1131 comprising amino acids 1 through 280 of Figure 230 (SEQ ID NO:319), the native sequence PRO1281 is a full-length or mature native sequence PRO1281 comprising amino acids 1 or about 16 to 775 of Figure 233 (SEQ ID NO:326), the native sequence PRO1064 is a mature or full-length native sequence PRO1064 comprising amino acids 1 to 153 of Figure 235 (SEQ ID NO:334), the native sequence PRO1379 is a full-length or mature native sequence PRO1379 comprising amino acids 1 or about 18 to 574 of Figure 238 (SEQ ID NO:340), the native sequence PRO844 is a full-length or mature native sequence PRO844 comprising amino acids 1 or 20 through 111 of Figure 240 (SEQ ID NO:344), the native sequence PRO848 is a full-length or mature native sequence PRO848 comprising amino acids 1 or 36 through 600 of Figure 242 (SEQ ID NO:347), the native sequence PRO1097 is a full-length or mature native sequence PRO1097 comprising amino acids 1 or 21 through 91 of Figure 244 (SEQ ID NO:349), the native sequence PRO1153 is a mature or full-length native sequence PRO1153 comprising amino acids 1 to 197 of Figure 246 (SEQ ID NO:351), the native sequence PRO1154 is a full-length or mature native sequence PRO1154 comprising amino acids 1 or 35 to 941 of Figure 248 (SEQ ID NO:353), the native sequence PRO1181 is a mature or full-length native sequence PRO1181 comprising amino acids 1 to 437 of Figure 250 (SEQ ID NO:355), the native sequence PRO1182 is a mature or full-length native sequence PRO1182 comprising amino acids 1 to 271 of Figure 252 (SEQ ID NO:357), the native sequence PRO1155 is a full-length native or mature sequence PRO1155 comprising amino acids 1 or 19 through 135 of Figure 254 (SEQ ID NO:359), the native sequence PRO1156 is a full-length or mature native sequence PRO1156 comprising amino acids 1 or about 23 to 159 of Figure 256 (SEQ ID NO:361), the native sequence PRO1098 is a full-length or mature native sequence PRO1098 comprising amino acids 1 or 20 through 78 of Figure 258 (SEQ ID NO:363), the native sequence PRO1127 is a full-length or mature native sequence PRO1127 comprising amino acids 1 or about 30 through 67 of Figure 260 (SEQ ID NO:365), the native sequence PRO1126 is a mature or full-length native sequence PRO1126 comprising amino acids 1 to 402 of Figure 262 (SEQ ID NO:367), the native sequence PRO1125 is a mature or full-length native sequence PRO1125 comprising amino acids 26 to 447 of Figure 264 (SEQ ID NO:369), the native sequence PRO1186 is a full-length or mature native sequence PRO1186 comprising amino acids 1 or about 20 through 105 of Figure 266 (SEQ ID NO:371), the native sequence PRO1198 is a full-length or mature native sequence PRO1198 comprising amino acids 1 or about 35 to 229 of Figure 268 (SEQ ID NO:373), the native sequence PRO1158 is a full-length or mature native sequence PRO1158 comprising amino acids 1 or about 20 to 123 of Figure 270 (SEQ ID NO:375), the native sequence PRO1159 is a mature or full-length native sequence PRO1159 comprising amino acids 1 to 90 of Figure 272 (SEQ ID NO:377), the native sequence PRO1124 is a mature or full-length native sequence PRO1124 comprising amino acids 22 through 919

of Figure 274 (SEQ ID NO:379), the native sequence PRO1287 is a mature or full-length native sequence PRO1287 comprising amino acids 1 to 532 of Figure 276 (SEQ ID NO:381), the native sequence PRO1312 is a full-length or mature native sequence PRO1312 comprising amino acids 1 or about 15 to 212 of Figure 278 (SEQ ID NO:387), the native sequence PRO1192 is a full-length or mature native sequence PRO1192 comprising amino acids 1 or about 22 to 215 of Figure 280 (SEQ ID NO:389), the native sequence PRO1160 is a mature or full-length native sequence PRO1160 comprising amino acids 1 to 90 of Figure 282 (SEQ ID NO:394), the native sequence PRO1187 is a full-length or mature native sequence PRO1187 comprising amino acids 1 or about 18 through 120 of Figure 284 (SEQ ID NO:399), the native sequence PRO1185 is a full-length or mature native sequence PRO1185 comprising amino acids 1 or about 22 through 198 of Figure 286 (SEQ ID NO:401), the native sequence PRO1345 is a mature or full-length native sequence PRO1345 comprising amino acids 1 to 206 of Figure 288 (SEQ ID NO:403), the native sequence PRO1245 is a full-length or mature native sequence PRO1245 comprising amino acids 1 or about 19 to 104 of Figure 290 (SEQ ID NO:408), the native sequence PRO1358 is a full-length or mature native sequence PRO1358 comprising amino acids 1 or about 19 through 444 of Figure 292 (SEQ ID NO:410), the native sequence PRO1195 is a full-length or mature native sequence PRO1195 comprising amino acids 1 or about 23 through 151 of Figure 294 (SEQ ID NO:412), the native sequence PRO1270 is a mature or full-length native sequence PRO1270 comprising amino acids 1 to 313 of Figure 296 (SEQ ID NO:414), the native sequence PRO1271 is a mature or full-length native sequence PRO1271 comprising amino acids 1 to 208 of Figure 298 (SEQ ID NO:416), the native sequence PRO1375 is a full-length or mature native sequence PRO1375 comprising amino acids 1 through 198 of Figure 300 (SEQ ID NO:418), the native sequence PRO1385 is a mature or full-length native sequence PRO1385 comprising amino acids 1 to 128 of Figure 302 (SEQ ID NO:420), the native sequence PRO1387 is a mature or full-length native sequence PRO1387 comprising amino acids 1 to 394 of Figure 304 (SEQ ID NO:422) and the native sequence PRO1384 is a full-length or mature native sequence PRO1384 comprising amino acids 1 to 229 of Figure 306 (SEQ ID NO:424). Start and stop codons are shown in bold font and underlined in the figures.

The PRO polypeptide "extracellular domain" or "ECD" refers to a form of the PRO polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a PRO polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the PRO polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified. Optionally, therefore, an extracellular domain of a PRO polypeptide may contain from about 5 or fewer amino acids on either or the transmembrane domain as initially identified.

"PRO polypeptide variant" means an active PRO polypeptide as defined above or below having at least about 80% amino acid sequence identity with a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Such PRO polypeptide variants include, for instance, PRO

polypeptides wherein one or more amino acid residues are added, or deleted, at the N- or C-terminus of the full-length native amino acid sequence. Ordinarily, a PRO polypeptide variant will have at least about 80% amino acid sequence identity, preferably at least about 81% amino acid sequence identity, more preferably at least about 82% amino acid sequence identity, more preferably at least about 83% amino acid sequence identity, more preferably at least about 84% amino acid sequence identity, more preferably at least about 85% amino acid sequence identity, more preferably at least about 86% amino acid sequence identity, more preferably at least about 87% amino acid sequence identity, more preferably at least about 88% amino acid sequence identity, more preferably at least about 89% amino acid sequence identity, more preferably at least about 90% amino acid sequence identity, more preferably at least about 91% amino acid sequence identity, more preferably at least about 92% amino acid sequence identity, more preferably at least about 93% amino acid sequence identity, more preferably at least about 94% amino acid sequence identity, more preferably at least about 95% amino acid sequence identity, more preferably at least about 96% amino acid sequence identity, more preferably at least about 97% amino acid sequence identity, more preferably at least about 98% amino acid sequence identity and most preferably at least about 99% amino acid sequence identity with the amino acid sequence of the full-length native amino acid sequence as disclosed herein. Ordinarily, PRO variant polypeptides are at least about 10 amino acids in length, often at least about 20 amino acids in length, more often at least about 30 amino acids in length, more often at least about 40 amino acids in length, more often at least about 50 amino acids in length, more often at least about 60 amino acids in length, more often at least about 70 amino acids in length, more often at least about 80 amino acids in length, more often at least about 90 amino acids in length, more often at least about 100 amino acids in length, more often at least about 150 amino acids in length, more often at least about 200 amino acids in length, more often at least about 300 amino acids in length, or more.

"Percent (%) amino acid sequence identity" with respect to the PRO polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific PRO polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the WU-BLAST-2 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. For purposes herein, a % amino acid sequence identity value is determined by dividing (a) the number of matching identical amino acid residues between the amino acid sequence of the PRO polypeptide of interest having a sequence derived from the native PRO polypeptide and the comparison amino acid sequence of interest (i.e., the sequence against which the PRO polypeptide of interest is being compared which may be

a PRO variant polypeptide) as determined by WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest.

"PRO variant polynucleotide" or "PRO variant nucleic acid sequence" means a nucleic acid molecule which encodes an active PRO polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, a PRO variant polynucleotide will have at least about 80% nucleic acid sequence identity, more preferably at least about 81% nucleic acid sequence identity, more preferably at least about 82% nucleic acid sequence identity, more preferably at least about 83% nucleic acid sequence identity, more preferably at least about 84% nucleic acid sequence identity, more preferably at least about 85% nucleic acid sequence identity, more preferably at least about 86% nucleic acid sequence identity, more preferably at least about 87% nucleic acid sequence identity, more preferably at least about 88% nucleic acid sequence identity, more preferably at least about 89% nucleic acid sequence identity, more preferably at least about 90% nucleic acid sequence identity, more preferably at least about 91% nucleic acid sequence identity, more preferably at least about 92% nucleic acid sequence identity, more preferably at least about 93% nucleic acid sequence identity, more preferably at least about 94% nucleic acid sequence identity, more preferably at least about 95% nucleic acid sequence identity, more preferably at least about 96% nucleic acid sequence identity, more preferably at least about 97% nucleic acid sequence identity, more preferably at least about 98% nucleic acid sequence identity and yet more preferably at least about 99% nucleic acid sequence identity with the nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

Ordinarily, PRO variant polynucleotides are at least about 30 nucleotides in length, often at least about 60 nucleotides in length, more often at least about 90 nucleotides in length, more often at least about 120 nucleotides in length, more often at least about 150 nucleotides in length, more often at least about 180 nucleotides in length, more often at least about 210 nucleotides in length, more often at least about 240 nucleotides in length, more often at least about 270 nucleotides in length, more often at least about 300 nucleotides in length, more often at least about 450 nucleotides in length, more often at least about 600 nucleotides in length, more often at least about 900 nucleotides in length, or more.

"Percent (%) nucleic acid sequence identity" with respect to PRO-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the PRO nucleic acid sequence of interest, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR)

software. For purposes herein, however, % nucleic acid sequence identity values are generated using the WU-BLAST-2 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. For purposes herein, a % nucleic acid sequence identity value is determined by dividing (a) the number of matching identical nucleotides between the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest having a sequence derived from the native sequence PRO polypeptide-encoding nucleic acid and the comparison nucleic acid molecule of interest (i.e., the sequence against which the PRO polypeptide-encoding nucleic acid molecule of interest is being compared which may be a variant PRO polynucleotide) as determined by WU-BLAST-2 by (b) the total number of nucleotides of the PRO polypeptide-encoding nucleic acid molecule of interest.

In other embodiments, PRO variant polynucleotides are nucleic acid molecules that encode an active PRO polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length PRO polypeptide as disclosed herein. PRO variant polypeptides may be those that are encoded by a PRO variant polynucleotide.

The term "positives", in the context of sequence comparison performed as described above, includes residues in the sequences compared that are not identical but have similar properties (e.g. as a result of conservative substitutions, see Table 1 below). For purposes herein, the % value of positives is determined by dividing (a) the number of amino acid residues scoring a positive value between the PRO polypeptide amino acid sequence of interest having a sequence derived from the native PRO polypeptide sequence and the comparison amino acid sequence of interest (i.e., the amino acid sequence against which the PRO polypeptide sequence is being compared) as determined in the BLOSUM62 matrix of WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest.

"Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one component of the PRO polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" PRO polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the PRO polypeptide nucleic acid. An isolated PRO polypeptide nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated PRO polypeptide nucleic acid molecules therefore are distinguished from the specific PRO polypeptide nucleic acid molecule as it exists in natural cells. However,

an isolated PRO polypeptide nucleic acid molecule includes PRO polypeptide nucleic acid molecules contained in cells that ordinarily express the PRO polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

The term "antibody" is used in the broadest sense and specifically covers, for example, single anti-PRO monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-PRO antibody compositions with polypeptidic specificity, single chain anti-PRO antibodies, and fragments of anti-PRO antibodies (see below). The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M

sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 μ g/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

"Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a PRO polypeptide fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site of an antibody (i.e., is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

"Active" or "activity" for the purposes herein refers to form(s) of a PRO polypeptide which retain a biological and/or an immunological activity of native or naturally-occurring PRO, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring PRO other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO and an "immunological" activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO.

The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native PRO polypeptide disclosed herein. In a similar manner, the term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native PRO polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically

include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native PRO polypeptides, peptides, small organic molecules, etc. Methods for identifying agonists or antagonists of a PRO polypeptide may comprise contacting a PRO polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the PRO polypeptide.

5 "Treatment" refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented.

"Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to
10 an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. "Intermittent" administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

"Mammal" for purposes of treatment refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats,
15 rabbits, etc. Preferably, the mammal is human.

Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often
20 the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including
25 glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, polyethylene glycol (PEG), and PLURONICS™.

"Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')₂, and Fv
30 fragments; diabodies; linear antibodies (Zapata et al., Protein Eng. 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an F(ab')₂ fragment that has two antigen-combining sites
35 and is still capable of cross-linking antigen.

"Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the VH-VL dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific
5 for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge
10 region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')₂ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant
15 domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2.

"Single-chain Fv" or "sFv" antibody fragments comprise the VH and VL domains of antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the VH and VL domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).
20

The term "diabodies" refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain (VH) connected to a light-chain variable domain (VL) in the same polypeptide chain (VH - VL). By using a linker that is too short to allow pairing between the two domains on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and
25 Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).
30

An "isolated" antibody is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1)
35 to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or

nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody in situ within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody so as to generate a "labeled" antibody. The label may be detectable by itself (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

By "solid phase" is meant a non-aqueous matrix to which the antibody of the present invention can adhere. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug (such as a PRO polypeptide or antibody thereto) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

A "small molecule" is defined herein to have a molecular weight below about 500 Daltons.

II. Compositions and Methods of the Invention

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO polypeptides. In particular, cDNAs encoding various PRO polypeptides have been identified and isolated, as disclosed in further detail in the Examples below. It is noted that proteins produced in separate expression rounds may be given different PRO numbers but the UNQ number is unique for any given DNA and the encoded protein, and will not be changed. However, for sake of simplicity, in the present specification the protein encoded by the full length native nucleic acid molecules disclosed herein as well as all further native homologues and variants included in the foregoing definition of PRO, will be referred to as "PRO/number", regardless of their origin or mode of preparation.

As disclosed in the Examples below, various cDNA clones have been deposited with the ATCC. The actual nucleotide sequences of those clones can readily be determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the PRO polypeptides and encoding nucleic acids described herein, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

1. Full-length PRO281 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO281 (UNQ244). In particular, cDNA encoding a PRO281

polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST-2 sequence alignment computer program, it has been found that a full-length native sequence PRO281 (shown in Figure 2 and SEQ ID NO:2) has certain amino acid sequence identity with the rat TEGT protein. Accordingly, it is presently believed that PRO281 disclosed in the present application is a newly identified TEGT homolog and may possess activity typical of that protein.

5

2. Full-length PRO276 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO276 (UNQ243). In particular, cDNA encoding a PRO276 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

10

As far as is known, the DNA16435-1208 sequence encodes a novel factor designated herein as PRO276; using WU-BLAST-2 sequence alignment computer programs, no significant sequence identities to any known proteins were revealed. The sequence identity identifications which were found are listed below in the examples.

3. Full-length PRO189 Polypeptides

15

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO189. In particular, Applicants have identified and isolated cDNA encoding a PRO189 polypeptide, as disclosed in further detail in the Examples below. To Applicants present knowledge, the DNA21624-1391 nucleotide sequence encodes a novel factor; using BLAST and FastA sequence alignment computer programs, no significant sequence identities to any known proteins were revealed.

20

4. Full-length PRO190 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO190. In particular, Applicants have identified and isolated cDNA encoding a PRO190 polypeptide, as disclosed in further detail in the Examples below. The PRO190-encoding clone was isolated from a human retina library. To Applicants present knowledge, the DNA23334-1392 nucleotide sequence encodes a novel multiple transmembrane spanning protein; using BLAST and FastA sequence alignment computer programs, there is some sequence identity with CMP-sialic acid and UDP-galactose transporters, indicating that PRO190 may be related to transporter or that PRO190 may be a novel transporter.

30

5. Full-length PRO341 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO341 (UNQ300). In particular, cDNA encoding a PRO341 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

35

The DNA26288-1239 clone was isolated from a human placenta library. As far as is known, the DNA26288-1239 sequence encodes a novel factor designated herein as PRO341; using the WU-BLAST-2 sequence alignment computer program, no significant sequence identities to any known proteins were revealed.

6. Full-length PRO180 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO180 (UNQ154). In particular, cDNA encoding a PRO180 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

The DNA26843-1389 clone was isolated from a human placenta library using oligos formed from DNA12922 isolated from an amylase screen. As far as is known, the DNA26843-1389 sequence encodes a novel factor designated herein as PRO180.

7. Full-length PRO194 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO194. In particular, Applicants have identified and isolated cDNA encoding a PRO194 polypeptide, as disclosed in further detail in the Examples below. The PRO194-encoding clone was isolated from a human fetal lung library. To Applicants present knowledge, the DNA26844-1394 nucleotide sequence encodes a novel factor; using BLAST and FastA sequence alignment computer programs, no significant sequence identities to any known proteins were revealed.

8. Full-length PRO203 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO203. In particular, Applicants have identified and isolated cDNA encoding a PRO203 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO203 polypeptide has sequence identity with GST ATPase. Accordingly, it is presently believed that PRO203 polypeptide disclosed in the present application is a newly identified member of the ATPase family and possesses activity typical of the GST ATPase.

9. Full-length PRO290 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO290. In particular, cDNA encoding a PRO290 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 23 (SEQ ID NO:33), revealed sequence identities between the PRO290 amino acid sequence and the following Dayhoff sequences: P_R99800, CC4H_HUMAN, YCS2_YEAST, CEF35G12_13, HSFAN_1, MMU52461_1, MMU70015_1, HSU67615_1, CET01H10_8 and CELT28F2_6.

It is currently believed that PRO290 is an intracellular protein related to one or more of the above proteins.

10. **Full-length PRO874 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO874. In particular, Applicants have identified and isolated cDNA encoding a PRO874 polypeptide, as disclosed in further detail in the Examples below. The PRO874-encoding clone was isolated from a human fetal lung library. To Applicants present knowledge, the DNA40621-1440 nucleotide sequence encodes a novel factor. Although, using BLAST and FastA sequence alignment computer programs, some sequence identity with known proteins was revealed.

11. **Full-length PRO710 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO710. In particular, Applicants have identified and isolated cDNA encoding a PRO710 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO710 polypeptide has significant similarity to the CDC45 protein. Accordingly, it is presently believed that PRO710 polypeptide disclosed in the present application is a newly identified CDC45 homolog.

12. **Full-length PRO1151 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1151. In particular, cDNA encoding a PRO1151 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST-2 sequence alignment computer program, it has been found that a full-length native sequence PRO1151 (shown in Figure 30 and SEQ ID NO:47) has certain amino acid sequence identity with the human 30 kD adipocyte complement-related precursor protein (ACR3_HUMAN). Accordingly, it is presently believed that PRO1151 disclosed in the present application is a newly identified member of the complement protein family and may possess activity typical of that family.

13. **Full-length PRO1282 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1282. In particular, cDNA encoding a PRO1282 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

As far as is known, the DNA45495-1550 sequence encodes a novel factor designated herein as PRO1282. Using WU-BLAST-2 sequence alignment computer programs, some sequence identities between PRO1282 and other leucine rich repeat proteins were revealed, as discussed in the examples below, indicating that a novel member of the leucine rich repeat superfamily has been identified.

14. **Full-length PRO358 Polypeptides**

The present invention further provides newly identified and isolated nucleotide sequences encoding a polypeptide referred to in the present application as PRO358. In particular, Applicants have identified and

isolated cDNA encoding a novel human Toll polypeptide (PRO358), as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the coding sequence of PRO358 shows significant homology to DNA sequences HSU88540_1, HSU88878_1, HSU88879_1, HSU88880_1, HS88881_1, and HSU79260_1 in the GenBank database. With the exception of HSU79260_1, the noted proteins have been identified as human toll-like receptors.

5 Accordingly, it is presently believed that the PRO358 proteins disclosed in the present application are newly identified human homologues of the *Drosophila* protein Toll, and are likely to play an important role in adaptive immunity. More specifically, PRO358 may be involved in inflammation, septic shock, and response to pathogens, and play possible roles in diverse medical conditions that are aggravated by immune response, such as, for example, diabetes, ALS, cancer, rheumatoid arthritis, and ulcers. The role of PRO385 as pathogen
10 pattern recognition receptors, sensing the presence of conserved molecular structures present on microbes, is further supported by the data disclosed in the present application, showing that a known human Toll-like receptor, TLR2 is a direct mediator of LPS signaling.

15. Full-length PRO1310 Polypeptides

15 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1310. In particular, cDNA encoding a PRO1310 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

 Using WU-BLAST-2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1310 (shown in Figure 36 and SEQ ID NO:62) has certain amino acid sequence identity with
20 carboxypeptidase X2. Accordingly, it is presently believed that PRO1310 disclosed in the present application is a newly identified member of the carboxypeptidase family and may possess carboxyl end amino acid removal activity.

16. Full-length PRO698 Polypeptides

25 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO698. In particular, Applicants have identified and isolated cDNA encoding a PRO698 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO698 polypeptide has significant similarity to the olfactomedin protein. Accordingly, it is presently believed that PRO698 polypeptide disclosed in the
30 present application may be a newly identified olfactomedin homolog.

17. Full-length PRO732 Polypeptides

 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO732. In particular, Applicants have identified and isolated cDNA
35 encoding a PRO732 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO732 polypeptide has significant similarity to the human placental Diff33 protein. Accordingly, it is presently believed that PRO732 polypeptide disclosed

in the present application is a newly identified Diff33 homolog.

18. **Full-length PRO1120 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1120. In particular, cDNA encoding a PRO1120 polypeptide has
5 been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST-2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1120 (shown in Figure 47 and SEQ ID NO:84) has certain amino acid sequence identity with the known sulfatase proteins designated CELK09C4_1, and GL6S_HUMAN, respectively, in the Dayhoff database (version 35.45 SwissProt 35). Accordingly, it is presently believed that PRO1120 disclosed in the present
10 application is a newly identified member of the sulfatase family and may possess activity typical of sulfatases.

19. **Full-length PRO537 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO537. In particular, cDNA encoding a PRO537 polypeptide has been
15 identified and isolated, as disclosed in further detail in the Examples below. The DNA49141-1431 clone was isolated from a human placenta library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the DNA49141-1431 clone does encode a secreted factor. As far as is known, the DNA49141-1431 sequence encodes a novel factor designated herein as PRO537; using the WU-BLAST2 sequence alignment computer program, no significant sequence identities to any known proteins were
20 revealed.

20. **Full-length PRO536 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO536. In particular, cDNA encoding a PRO536 polypeptide has been
25 identified and isolated, as disclosed in further detail in the Examples below.

The DNA49142-1430 clone was isolated from a human infant brain library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the DNA49142-1430 clone does encode a secreted factor. As far as is known, the DNA49142-1430 sequence encodes a novel factor designated herein as PRO536; using the WU-BLAST-2 sequence alignment computer program, no significant sequence
30 identities to any known proteins were revealed.

21. **Full-length PRO535 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO535. In particular, cDNA encoding a PRO535 polypeptide has been
35 identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO535 (shown in Figure 53 and SEQ ID NO:99) has amino acid sequence identity with a

putative peptidyl-prolyl isomerase protein. Accordingly, it is presently believed that PRO535 disclosed in the present application is a newly identified member of the isomerase protein family and may possess activity typical of those proteins.

22. **Full-length PRO718 Polypeptides**

5 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO718. In particular, Applicants have identified and isolated cDNA encoding a PRO718 polypeptide, as disclosed in further detail in the Examples below. The PRO718-encoding clone was isolated from a human fetal lung library. To Applicants present knowledge, the DNA49647-1398 nucleotide sequence encodes a novel factor; using BLAST and FastA sequence alignment computer programs,
10 no significant sequence identities to any known proteins were revealed.

23. **Full-length PRO872 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO872. In particular, Applicants have identified and isolated cDNA
15 encoding a PRO872 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO872 polypeptide has sequence identity with dehydrogenases. Accordingly, it is presently believed that PRO872 polypeptide disclosed in the present application is a newly identified member of the dehydrogenase family and possesses dehydrogenase activity.

20 24. **Full-length PRO1063 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1063. In particular, Applicants have identified and isolated cDNA encoding a PRO1063 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO1063 polypeptide has significant
25 similarity to the human type IV collagenase protein. Accordingly, it is presently believed that PRO1063 polypeptide disclosed in the present application is a newly identified collagenase homolog.

25. **Full-length PRO619 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO619. In particular, cDNA encoding a PRO619 polypeptide has been
30 identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST-2 sequence alignment computer program, it has been found that a full-length native sequence PRO619 (shown in Figure 68 and SEQ ID NO:117) has certain amino acid sequence identity with VpreB3. Accordingly, it is presently believed that PRO619 disclosed in the present application is a newly
35 identified member of the IgG superfamily and may possess activity related to the assembly and/or components of the surrogate light chain associated with developing B cells.

26. **Full-length PRO943 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO943. In particular, cDNA encoding a PRO943 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

5 Using the WU-BLAST-2 sequence alignment computer program, it has been found that a full-length native sequence PRO943 (shown in Figure 70 and SEQ ID NO:119) has amino acid sequence identity with the fibroblast growth factor receptor-4 protein. Accordingly, it is presently believed that PRO943 disclosed in the present application is a newly identified member of the fibroblast growth factor receptor family and may possess activity typical of that family.

10 27. **Full-length PRO1188 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1188. In particular, cDNA encoding a PRO1188 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

15 As discussed in more detail in Example 1 below, using WU-BLAST-2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1188 (shown in Figure 72; SEQ ID NO:124) has certain amino acid sequence identity with nucleotide pyrophosphohydrolase (SSU83114_1). Accordingly, it is presently believed that PRO1188 disclosed in the present application is a newly identified member of the nucleotide pyrophosphohydrolase family and may possess activity typical of that family of proteins.

20 28. **Full-length PRO1133 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1133. In particular, cDNA encoding a PRO1133 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

25 Using WU-BLAST-2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1133 (shown in Figure 74 and SEQ ID NO:129) has certain amino acid sequence identity with netrin 1a, Dayhoff accession AF002717_1. Accordingly, it is presently believed that PRO1133 disclosed in the present application shares at least one related mechanism with netrin.

30 29. **Full-length PRO784 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO784. In particular, cDNA, designated herein as "DNA53978-1443", which encodes a PRO784 polypeptide, has been identified and isolated, as disclosed in further detail in the Examples below.

35 Using BLAST and FastA sequence alignment computer programs, it has been found that a full-length native sequence PRO784 (shown in Figure 76 and SEQ ID NO:135) has certain amino acid sequence identity with sec22 homologs. Accordingly, it is presently believed that PRO784 disclosed in the present application is a newly identified member of the sec22 family and may possess vesicle trafficking activities typical of the sec22

family.

30. **Full-length PRO783 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO783. In particular, Applicants have identified and isolated cDNA
5 encoding a PRO783 polypeptide, as disclosed in further detail in the Examples below. The PRO783-encoding clone was isolated from a human fetal kidney library. To Applicants present knowledge, the DNA53996-1442 nucleotide sequence encodes a novel factor. However, using BLAST and FastA sequence alignment computer programs, some sequence identity to known proteins was found.

10 31. **Full-length PRO820 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO820. In particular, Applicants have identified and isolated cDNA encoding a PRO820 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA
15 sequence alignment computer programs, Applicants found that various portions of the PRO820 polypeptide have sequence identity with the low affinity immunoglobulin gamma Fc receptor, the IgE high affinity Fc receptor and the high affinity immunoglobulin epsilon receptor. Accordingly, it is presently believed that PRO820 polypeptide disclosed in the present application is a newly identified member of the Fc receptor family.

20 32. **Full-length PRO1080 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1080. In particular, Applicants have identified and isolated cDNA encoding a PRO1080 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and
25 FastA sequence alignment computer programs, Dayhoff database (version 35.45 SwissProt 35), Applicants found that the PRO1080 polypeptide has sequence identity with a 39.9 kd protein designated as "YRY1_CAEEL", a DnaJ homolog designated "AF027149_5", a DnaJ homolog 2 designated "RNU95727_1", and Dna3/Cpr3 designated "AF011793_1". Accordingly, these results indicate that the PRO1080 polypeptide disclosed in the present application may be a newly identified member of the DnaJ-like protein family and therefore may be involved in protein biogenesis.

30 33. **Full-length PRO1079 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1079. In particular, cDNA encoding a PRO1079 polypeptide has
been identified and isolated, as disclosed in further detail in the Examples below.

35 As far as is known, the DNA56050-1455 sequence encodes a novel factor designated herein as PRO1079. Although, using WU-BLAST2 sequence alignment computer programs, some sequence identities to known proteins was revealed.

34. Full-length PRO793 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO793. In particular, cDNA encoding a PRO793 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

5 The DNA56110-1437 clone was isolated from a human skin tumor library. As far as is known, the DNA56110-1437 sequence encodes a novel factor designated herein as PRO793; using the WU-BLAST-2 sequence alignment computer program, no significant sequence identities to any known proteins were revealed.

35. Full-length PRO1016 Polypeptides

10 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1016. In particular, Applicants have identified and isolated cDNA encoding a PRO1016 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that various portions of the PRO1016 polypeptide have sequence identity with acyltransferases. Accordingly, it is presently believed that PRO1016 polypeptide disclosed in the present application is a newly identified member of the acyltransferase family and
15 possesses acylation capabilities typical of this family.

36. Full-length PRO1013 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1013. In particular, Applicants have identified cDNA encoding a
20 PRO1013 polypeptide, as disclosed in further detail in the Examples below. The PRO1013-encoding clone came from a human breast tumor tissue library. Thus, the PRO1013-encoding clone may encode a secreted factor related to cancer. To Applicants present knowledge, the DNA56410-1414 nucleotide sequence encodes a novel factor. Using BLAST and FastA sequence alignment computer programs, some sequence identity with KIAA0157 and P120 was revealed. PRO1013 has at least one region in common with growth factor and
25 cytokine receptors.

37. Full-length PRO937 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO937. In particular, Applicants have identified and isolated cDNA
30 encoding a PRO937 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO937 polypeptide has significant sequence identity with members of the glypican family of proteins. Accordingly, it is presently believed that PRO937 polypeptide disclosed in the present application is a newly identified member of the glypican family possesses properties typical of the glypican family.

35

38. Full-length PRO842 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO842. In particular, cDNA encoding a PRO842 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

As far as is known, the DNA56855-1447 sequence encodes a novel secreted factor designated herein as PRO842. However, using WU-BLAST2 sequence alignment computer programs, some sequence identity to any known proteins were revealed.

39. Full-length PRO839 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO839. In particular, cDNA encoding a PRO839 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

As far as is known, the DNA56859-1445 sequence encodes a novel factor designated herein as PRO839. However, using WU-BLAST-2 sequence alignment computer programs, some sequence identities to known proteins was revealed.

40. Full-length PRO1180 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1180. In particular, Applicants have identified and isolated cDNA encoding a PRO1180 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO1180 polypeptide has significant similarity to methyltransferase enzymes. Accordingly, it is presently believed that PRO1180 polypeptide disclosed in the present application is a newly identified member of the methyltransferase family and possesses activity typical of that family.

41. Full-length PRO1134 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1134. In particular, cDNA encoding a PRO1134 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

The DNA56865-1491 clone was isolated from a human fetal liver spleen library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the DNA56865-1491 clone does encode a secreted factor. As far as is known, the DNA56865-1491 sequence encodes a novel factor designated herein as PRO1134; using the WU-BLAST2 sequence alignment computer program, no significant sequence identities to any known proteins were revealed.

42. Full-length PRO830 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO830. In particular, cDNA encoding a PRO830 polypeptide has been

identified and isolated, as disclosed in further detail in the Examples below.

The DNA56866-1342 clone was isolated from a human fetal liver/spleen library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the DNA56866-1342 clone does encode a secreted factor. As far as is known, the DNA56866-1342 sequence encodes a novel factor designated herein as PRO830; using the WU-BLAST-2 sequence alignment computer program, no significant
5 sequence identities to any known proteins were revealed.

43. Full-length PRO1115 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1115. In particular, cDNA encoding a PRO1115 polypeptide has
10 been identified and isolated, as disclosed in further detail in the Examples below.

As far as is known, the DNA56868-1478 sequence encodes a novel transmembrane protein designated herein as PRO1115. Although, using WU-BLAST-2 sequence alignment computer programs, some sequence identities to known proteins were revealed.

15 44. Full-length PRO1277 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1277. In particular, cDNA encoding a PRO1277 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST-2 sequence alignment computer programs, it has been found that a full-length native
20 sequence PRO1277 (shown in Figure 113 and SEQ ID NO:179) has certain amino acid sequence identity with Coch-5B2 protein (designated "AF012252_1" in the Dayhoff database). Accordingly, it is presently believed that PRO1277 disclosed in the present application is a newly identified member of the Coch-5B2 protein family and may possess the same activities and properties as Coch-5B2.

25 45. Full-length PRO1135 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1135. In particular, Applicants have identified and isolated cDNA encoding a PRO1135 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and
30 FastA sequence alignment computer programs, Applicants found that the PRO1135 polypeptide has significant similarity to the alpha 1,2-mannosidase protein. Accordingly, it is presently believed that PRO1135 polypeptide disclosed in the present application is a newly identified member of the mannosidase enzyme family and possesses activity typical of that family of proteins.

35 46. Full-length PRO1114 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1114 interferon receptor. In particular, cDNA encoding a PRO1114
interferon receptor polypeptide has been identified and isolated, as disclosed in further detail in the Examples

below.

Using the WU-BLAST-2 sequence alignment computer program, it has been found that a full-length native sequence PRO1114 interferon receptor polypeptide (shown in Figure 117 and SEQ ID NO:183) has sequence identity with the other known interferon receptors. Accordingly, it is presently believed that PRO1114 interferon receptor possesses activity typical of other interferon receptors.

5

47. Full-length PRO828 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO828. In particular, Applicants have identified and isolated cDNA encoding a PRO828 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA
10 sequence alignment computer programs, Applicants found that the PRO828 polypeptide has sequence identity with glutathione peroxidases. Accordingly, it is presently believed that PRO828 polypeptide disclosed in the present application is a newly identified member of the glutathione peroxidase family and possesses peroxidase activity and other properties typical of glutathione peroxidases.

15

48. Full-length PRO1009 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1009. In particular, cDNA encoding a PRO1009 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST-2 sequence alignment computer programs, it has been found that a full-length native
20 sequence PRO1009 (shown in Figure 122 and SEQ ID NO:194) has certain amino acid sequence identity with long-chain acyl-CoA synthetase homolog designated "F69893". Accordingly, it is presently believed that PRO1009 disclosed in the present application is a newly identified member of the long-chain acyl-CoA synthetase family and may possess activity related to this family.

25

49. Full-length PRO1007 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1007. In particular, Applicants have identified and isolated cDNA encoding a PRO1007 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and
30 FastA sequence alignment computer programs, Applicants found that various portions of the PRO1007 polypeptide have sequence identity with MAGPIAP. Accordingly, it is presently believed that PRO1007 polypeptide disclosed in the present application is a newly identified member of the MAGPIAP family and is associated with metastasis and/or cell signaling and/or cell replication.

50. Full-length PRO1056 Polypeptides

35 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1056. In particular, cDNA encoding a PRO1056 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST-2 sequence alignment computer program, it has been found that a full-length native sequence PRO1056 (shown in Figure 127 and SEQ ID NO:199) has amino acid sequence identity with a chloride channel protein. Accordingly, it is presently believed that PRO1056 disclosed in the present application is a newly identified chloride channel protein homolog.

5 **51. Full-length PRO826 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO826. In particular, cDNA encoding a PRO826 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

10 The DNA57694-1341 clone was isolated from a human fetal heart library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the DNA57694-1341 clone does encode a secreted factor. As far as is known, the DNA57694-1341 sequence encodes a novel factor designated herein as PRO826; using the WU-BLAST-2 sequence alignment computer program, no significant sequence identities to any known proteins were revealed.

15 **52. Full-length PRO819 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO819. In particular, cDNA encoding a PRO819 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

20 The DNA57695-1340 clone was isolated from a human fetal liver spleen library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the DNA57695-1340 clone does encode a secreted factor. As far as is known, the DNA57695-1340 sequence encodes a novel factor designated herein as PRO819; using the WU-BLAST-2 sequence alignment computer program, no significant sequence identities to any known proteins were revealed.

25 **53. Full-length PRO1006 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1006. In particular, Applicants have identified and isolated cDNA encoding a PRO1006 polypeptide, as disclosed in further detail in the Examples below. The PRO1006-encoding clone was isolated from a human uterus library. To Applicants present knowledge, the DNA57699-1412 nucleotide sequence encodes a novel factor; using BLAST and FastA sequence alignment computer programs, some sequence identity with a putative tyrosine protein kinase was revealed.

54. Full-length PRO1112 Polypeptides

35 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1112. In particular, Applicants have identified cDNA encoding a PRO1112 polypeptide, as disclosed in further detail in Example 1 below. To Applicants present knowledge, the DNA57702-1476 nucleotide sequence encodes a novel factor, although using BLAST and FastA sequence

alignment computer programs some sequence identity with other known proteins was found.

55. **Full-length PRO1074 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1074. In particular, Applicants have identified and isolated cDNA
5 encoding a PRO1074 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO1074 polypeptide has sequence identity with galactosyltransferase. Accordingly, it is presently believed that PRO1074 polypeptide disclosed in the present application is a newly identified member of the galactosyltransferase family and possesses galactosyltransferase activity.

56. **Full-length PRO1005 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1005. In particular, cDNA encoding a PRO1005 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

15 As far as is known, the DNA57708-1411 sequence encodes a novel factor designated herein as PRO1005. However, using WU-BLAST2 sequence alignment computer programs, some sequence identities with known proteins was revealed.

57. **Full-length PRO1073 Polypeptides**

20 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1073. In particular, cDNA encoding a PRO1073 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

As far as is known, the DNA57710 sequence encodes a novel secreted factor designated herein as PRO1073. However, using WU-BLAST2 sequence alignment computer programs, some sequence identities to
25 known proteins were revealed.

58. **Full-length PRO1152 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1152. In particular, cDNA encoding a PRO1152 polypeptide has
30 been identified and isolated, as disclosed in further detail in the Examples below.

The DNA57711-1501 clone was isolated from a human infant brain library. As far as is known, the DNA57711-1501 sequence encodes a novel factor designated herein as PRO1152; using the WU-BLAST-2 sequence alignment computer program, no significant sequence identities to any known proteins were revealed.

59. **Full-length PRO1136 Polypeptides**

35 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1136. In particular, cDNA encoding a PRO1136 polypeptide has

been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1136 (shown in Figure 147 and SEQ ID NO:219) has amino acid sequence identity with PDZ domain-containing proteins. Accordingly, it is presently believed that PRO1136 disclosed in the present application is a newly identified member of the PDZ domain-containing protein family and may possess activity typical of that family.

60. **Full-length PRO813 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO813. In particular, Applicants have identified and isolated cDNA encoding a PRO813 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO813 polypeptide has significant similarity to the pulmonary surfactant-associated protein C. Accordingly, it is presently believed that PRO813 polypeptide disclosed in the present application is a newly identified pulmonary surfactant-associated protein C homolog.

61. **Full-length PRO809 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO809. In particular, Applicants have identified and isolated cDNA encoding a PRO809 polypeptide, as disclosed in further detail in the Examples below. To Applicants present knowledge, the DNA57836-1338 nucleotide sequence encodes a novel factor.

62. **Full-length PRO791 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO791. In particular, Applicants have identified and isolated cDNA encoding a PRO791 polypeptide, as disclosed in further detail in the Examples below. To Applicants present knowledge, the DNA57838-1337 nucleotide sequence encodes a novel factor; however, using BLAST and FastA sequence alignment computer programs, there does appear to be some sequence identity with MHC-I antigens, indicating that PRO791 may be related thereto in structure and function.

63. **Full-length PRO1004 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1004. In particular, cDNA encoding a PRO1004 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

As far as is known, the DNA57844-1410 sequence encodes a novel factor designated herein as PRO1004. However, using WU-BLAST2 sequence alignment computer programs, some sequence identities with known proteins were revealed.

64. Full-length PRO1111 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1111. In particular, cDNA encoding a PRO1111 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1111 (shown in Figure 157 and SEQ ID NO:229) has certain amino acid sequence identity with
5 LIG. Accordingly, it is presently believed that PRO1111 disclosed in the present application is a newly identified member of this glycoprotein family.

65. Full-length PRO1344 Polypeptides

10 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1344. In particular, cDNA encoding a PRO1344 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1344 (shown in Figure 159 and SEQ ID NO:231) has certain amino acid sequence identity
15 with the factor C protein of Carcinoscopus rotundicauda. Accordingly, it is presently believed that PRO1344 disclosed in the present application is a newly identified factor C protein and may possess activity typical of that protein.

66. Full-length PRO1109 Polypeptides

20 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1109. In particular, cDNA encoding a PRO1109 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1109 (shown in Figure 161 and SEQ ID NO:236) has certain amino acid sequence identity
25 with the human UDP-Gal:GlcNAc galactosyltransferase protein. Accordingly, it is presently believed that PRO1109 disclosed in the present application is a newly identified β -galactosyltransferase enzyme and has activity typical of those enzymes.

67. Full-length PRO1383 Polypeptides

30 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1383. In particular, cDNA encoding a PRO1383 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1383 (shown in Figure 163 and SEQ ID NO:241) has certain amino acid sequence identity
35 with the putative human transmembrane protein nmb precursor (NMB_HUMAN). Accordingly, it is presently believed that PRO1383 disclosed in the present application is a newly identified nmb homolog.

68. Full-length PRO1003 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1003. In particular, Applicants have identified and isolated cDNA encoding a PRO1003 polypeptide, as disclosed in further detail in the Examples below. The PRO1003-encoding clone was isolated from a human breast tumor tissue library. The PRO1003-encoding clone was isolated using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the PRO1003-encoding clone may encode a secreted factor. To Applicants present knowledge, the UNQ487 (DNA58846-1409) nucleotide sequence encodes a novel factor; using BLAST and FastA sequence alignment computer programs, no sequence identities to any known proteins were revealed.

69. Full-length PRO1108 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1108. In particular, Applicants have identified and isolated cDNA encoding a PRO1108 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO1108 polypeptide has significant similarity to the LPAAT protein. Accordingly, it is presently believed that PRO1108 polypeptide disclosed in the present application is a newly identified LPAAT homolog.

70. Full-length PRO1137 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1137. In particular, Applicants have identified and isolated cDNA encoding a PRO1137 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO1137 polypeptide has sequence identity with ribosyltransferases. Accordingly, it is presently believed that PRO1137 polypeptide disclosed in the present application is a newly identified member of the ribosyltransferase family and possesses ribosyltransferase activity.

71. Full-length PRO1138 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1138. In particular, Applicants have identified and isolated cDNA encoding a PRO1138 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO1138 polypeptide has sequence identity with CD84 leukocyte antigen. Accordingly, it is presently believed that PRO1138 polypeptide disclosed in the present application is a newly identified member of the Ig superfamily and has activity typical of other members of the Ig superfamily.

72. Full-length PRO1054 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1054. In particular, cDNA encoding a PRO1054 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1054 (shown in Figure 174 and SEQ ID NO:256) has amino acid sequence identity with one or more of the major urinary proteins. Accordingly, it is presently believed that PRO1054 disclosed in the present application is a newly identified member of the MUP family and may possess activity typical of that family.

73. Full-length PRO994 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO994. In particular, cDNA encoding a PRO994 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO994 (shown in Figure 176 and SEQ ID NO:258) has amino acid sequence identity with the tumor-associated antigen L6. Accordingly, it is presently believed that PRO994 disclosed in the present application is a newly identified L6 antigen homolog.

74. Full-length PRO812 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO812. In particular, cDNA encoding a PRO812 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO812 (shown in Figure 178 and SEQ ID NO:260) has amino acid sequence identity with the prostatic steroid-binding c1 protein. Accordingly, it is presently believed that PRO812 disclosed in the present application is a newly identified prostatic steroid-binding c1 protein homolog.

75. Full-length PRO1069 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1069. In particular, Applicants have identified and isolated cDNA encoding a PRO1069 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, it was found that the PRO1069 polypeptide has sequence identity with CHIF. Accordingly, it is presently believed that PRO1069 polypeptide disclosed in the present application is a newly identified CHIF polypeptide and is involved in ion conductance or regulation of ion conductance.

76. Full-length PRO1129 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1129. In particular, Applicants have identified and isolated cDNA encoding a PRO1129 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO1129 polypeptide has significant
5 similarity to the cytochrome P-450 family of proteins. Accordingly, it is presently believed that PRO1129 polypeptide disclosed in the present application is a newly identified member of the cytochrome P-450 family and possesses activity typical of that family.

77. Full-length PRO1068 Polypeptides

10 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1068. In particular, cDNA encoding a PRO1068 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1068 has amino acid sequence identity with urotensin. Accordingly, it is presently believed that
15 PRO1068 disclosed in the present application is a newly identified member of the urotensin family and may possess activity typical of the urotensin family.

78. Full-length PRO1066 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1066. In particular, Applicants have identified and isolated cDNA encoding a PRO1066 polypeptide, as disclosed in further detail in the Examples below. The PRO1066-encoding
20 clone was isolated from a human pancreatic tumor tissue library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the PRO1066-encoding clone may encode a secreted factor. To Applicants present knowledge, the DNA59215-1425 nucleotide sequence encodes a novel factor; using BLAST and FastA sequence alignment computer programs, no sequence identities to any known proteins
25 were revealed.

79. Full-length PRO1184 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1184. In particular, Applicants have identified cDNA encoding a
30 PRO1184 polypeptide, as disclosed in further detail in the Examples below. To Applicants present knowledge, the DNA59220-1514 nucleotide sequence encodes a novel secreted factor.

80. Full-length PRO1360 Polypeptides

35 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1360. In particular, cDNA encoding a PRO1360 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

As far as is known, the DNA59488-1603 sequence encodes a novel factor designated herein as PRO1360; using WU-BLAST2 sequence alignment computer programs, no significant sequence identities to any known proteins were revealed. Some sequence identities were revealed, as indicated below in the examples.

81. Full-length PRO1029 Polypeptides

5 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1029. In particular, cDNA encoding a PRO1029 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

10 The DNA59493-1420 clone was isolated from a human fetal liver spleen library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the DNA59493-1420 clone does encode a secreted factor. As far as is known, the DNA59493-1420 sequence encodes a novel factor designated herein as PRO1029; using the WU-BLAST2 sequence alignment computer program, no sequence identities to any known proteins were revealed.

82. Full-length PRO1139 Polypeptides

15 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1139. In particular, Applicants have identified and isolated cDNAs encoding PRO1139, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the human PRO1139 protein originally identified exhibits a significant sequence homology to the a OB receptor associated protein HSOBRGRP_1, described by Bailleul
20 et al., Nucleic Acids Res. 25, 2752-2758 (1997) (EMBL Accession No: Y12670).

83. Full-length PRO1309 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1309. In particular, cDNA encoding a PRO1309 polypeptide has
25 been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1309 (shown in Figure 196 and SEQ ID NO:278) has certain amino acid sequence identity with a protein designated KIAA0416, given the Dayhoff designation AB007876_1. Moreover, PRO1309 has leucine rich repeats, accordingly, it is presently believed that PRO1309 disclosed in the present application is
30 a newly identified member of the leucine rich protein family and may be involved in protein protein interactions.

84. Full-length PRO1028 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1028. In particular, Applicants have identified and isolated cDNA
35 encoding a PRO1028 polypeptide, as disclosed in further detail in the Examples below. To Applicants present knowledge, the DNA59603-1419 nucleotide sequence encodes a novel factor. BLAST and FastA sequence alignment computer programs showed some sequence identity with proteins such as those designated "A53050"

85. Full-length PRO1027 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1027. In particular, Applicants have identified and isolated cDNA
5 encoding a PRO1027 polypeptide, as disclosed in further detail in the Examples below. The PRO1027-encoding clone was identified in a human uterine cervical tissue library. To Applicants present knowledge, the DNA59605-1418 nucleotide sequence encodes a novel factor.

86. Full-length PRO1107 Polypeptides

10 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1107. In particular, Applicants have identified and isolated cDNA encoding a PRO1107 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO1107 polypeptide has some similarity to the PC-1 protein, human insulin receptor tyrosine kinase inhibitor, an alkaline phosphodiesterase,
15 and autotaxin. Accordingly, it is presently believed that PRO1107 polypeptide disclosed in the present application is a newly identified member of the phosphodiesterase family.

87. Full-length PRO1140 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding novel multi-
20 span transmembrane polypeptides referred to in the present application as PRO1140. In particular, Applicants have identified and isolated cDNA encoding a PRO1140 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, some sequence identity with known proteins was found.

88. Full-length PRO1106 Polypeptides

25 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1106. In particular, Applicants have identified and isolated cDNA encoding a PRO1106 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO1106 polypeptide has significant
30 similarity to the peroxisomal calcium-dependent solute carrier. Accordingly, it is presently believed that PRO1106 polypeptide disclosed in the present application is a newly identified member of the mitochondrial carrier superfamily and possesses transporter activity typical of this family.

89. Full-length PRO1291 Polypeptides

35 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1291. In particular, cDNA encoding a PRO1291 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1291 (shown in Figure 208 and SEQ ID NO:291) has certain amino acid sequence identity with the butyrophilin protein. Accordingly, it is presently believed that PRO1291 disclosed in the present application is a newly identified butyrophilin homolog and may possess activity typical of that protein.

5 **90. Full-length PRO1105 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1105. In particular, Applicants have identified cDNA encoding a PRO1105 polypeptide, as disclosed in further detail in the Examples below. To Applicants present knowledge, the DNA59612-1466 nucleotide sequence encodes a novel factor. There is, however, some sequence identity
10 with a peroxidase precursor designated in a Dayhoff database as "ATTS1623_1".

91. Full-length PRO511 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO511. In particular, Applicants have identified and isolated cDNA
15 encoding a PRO511 polypeptide, as disclosed in further detail in the Examples below. The PRO511-encoding clone was isolated from a human colon tissue library. To Applicants present knowledge, the DNA59613-1417 nucleotide sequence encodes a novel factor; using BLAST and FastA sequence alignment computer programs, sequence identities with RoBo-1, phospholipase inhibitors and a protein designated as "SSC20F10_1" were revealed, indicated that PRO511 may be related to one or more of these proteins.

20 **92. Full-length PRO1104 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1104. In particular, Applicants have identified and isolated cDNA encoding a PRO1104 polypeptide, as disclosed in further detail in the Examples below. To Applicants present
25 knowledge, the DNA59616-1465 nucleotide sequence encodes a novel factor; using BLAST and FastA sequence alignment computer programs, some sequence identity appeared with proteins designated as "AB002107_1", "AF022991_1" and "SP96_DICDI".

93. Full-length PRO1100 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1100. In particular, Applicants have identified cDNA encoding a PRO1100 polypeptide, as disclosed in further detail in the Examples below. To Applicants present knowledge, the DNA59619-1464 nucleotide sequence encodes a novel factor; using BLAST and FastA sequence alignment
30 computer programs, only some sequence identity with known proteins was revealed. There is some sequence
35 identity with the yeast hypothetical 42.5 KD protein in TSM1-ARE1 intergenic region (ACCESSION NO:140496), designated "YSCT4_YEAST".

94. **Full-length PRO836 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO836. In particular, Applicants have identified and isolated cDNA encoding a PRO836 polypeptide, as disclosed in further detail in the Examples below. To Applicants present knowledge, the DNA59620-1463 nucleotide sequence encodes a novel factor. Using BLAST and FastA
5 sequence alignment computer programs, there appears to be some sequence identity with SLS1.

95. **Full-length PRO1141 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1141. In particular, cDNA encoding a PRO1141 polypeptide has
10 been identified and isolated, as disclosed in further detail in the Examples below.

The DNA59625-1498 clone was isolated from a human ileum tissue library. As far as is known, the DNA59625-1498 sequence encodes a novel factor designated herein as PRO1141; using the WU-BLAST2 sequence alignment computer program, no sequence identities to any known proteins were revealed.

15 96. **Full-length PRO1132 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1132. In particular, cDNA encoding a PRO1132 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST2 sequence alignment computer program, it has been found that a full-length native
20 sequence PRO1132 (shown in Figure 226 and SEQ ID NO:309) has certain amino acid sequence identity with enamel matrix serine proteinase 1 and neuropsin. Accordingly, it is presently believed that PRO1132 disclosed in the present application is a newly identified member of the serine protease family and may possess protease activity typical of this family.

25 97. **Full-length PRO1346 Polypeptides**

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as NL7 (UNQ701). In particular, cDNA encoding an NL7 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

As disclosed in the Examples below, a clone DNA59776-1600 has been deposited with ATCC. The
30 actual nucleotide sequence of the clone can be readily determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the NL7 (PRO1346) herein, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time of filing.

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native
35 sequence NL7 (shown in Figure 228 and SEQ ID NO:314) has certain amino acid sequence identity with microfibril-associated glycoprotein 4 (MFA4_HUMAN); ficolin-A - Mus musculus (AB007813_1); human lectin P35 (D63155S6_1); ficolin B - Mus musculus (AF0063217_1); human tenascin-R (restriction) (HS518E13_1);

the long form of a rat janusin precursor (A45445); fibrinogen-related protein HFREP-1 precursor (JNO596); a human Tenascin precursor (TENA HUMAN); human CDT6 (HSY16132_1); and angiopoietin-1 - Mus musculus (MMU83509_1). It is presently believed that NL7 disclosed in the present application is a novel TIE ligand homologue, and may play a role in angiogenesis and/or vascular maintenance and/or wound healing and/or inflammation and/or tumor development and/or growth

5

98. Full-length PRO1131 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1131. In particular, cDNA encoding a PRO1131 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

10

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1131 (shown in Figure 230 and SEQ ID NO:319) has certain amino acid sequence identity with a lectin-like oxidized LDL receptor. Accordingly, it is presently believed that PRO1131 disclosed in the present application may have at least one mechanism similar to those of the LDL receptors.

15

99. Full-length PRO1281 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1281. In particular, cDNA encoding a PRO1281 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

20

The DNA59820-1549 clone was isolated from a human fetal liver library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, as far as is known, the DNA59820-1549 sequence encodes a novel factor designated herein as PRO1281. Using WU-BLAST2 sequence alignment computer programs, some sequence identities to known proteins was found, but determined not to be significant.

25

100. Full-length PRO1064 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1064. In particular, cDNA encoding a PRO1064 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

30

The DNA59827-1426 clone was isolated from a human fetal kidney library. As far as is known, the DNA59827-1426 sequence encodes a novel factor designated herein as PRO1064; using the WU-BLAST2 sequence alignment computer program, no significant sequence identities to any known proteins were revealed.

101. Full-length PRO1379 Polypeptides

35

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1379. In particular, cDNA encoding a PRO1379 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

The DNA59828 clone was isolated from a human fetal kidney library. As far as is known, the PRO1379 polypeptide encoded thereby is a novel secreted factor. Using WU-BLAST2 sequence alignment

computer programs, sequence identity was found between PRO1379 and a hypothetical yeast protein "YHY8_YEAST" (Dayhoff database; version 35.45 SwissProt 35), particularly at the C-terminal ends. Sequence homologies with other known proteins were revealed, but determined not to be significant.

102. Full-length PRO844 Polypeptides

5 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO844. In particular, Applicants have identified and isolated cDNA encoding a PRO844 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO844 polypeptide has sequence identity with serine protease inhibitors. Accordingly, it is presently believed that PRO844 polypeptide disclosed in the
10 present application is a newly identified serine protease inhibitor and is capable of inhibiting serine proteases.

103. Full-length PRO848 Polypeptides

 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO848. In particular, Applicants have identified and isolated cDNA
15 encoding a PRO848 polypeptide, as disclosed in further detail in the Examples below. Using BLAST and FastA sequence alignment computer programs, Applicants found that the PRO848 polypeptide has sequence identity with sialyltransferases. Accordingly, it is presently believed that PRO848 polypeptide disclosed in the present application is a newly identified member of the sialyltransferase family and possesses sialylation capabilities as typical of this family.

20

104. Full-length PRO1097 Polypeptides

 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1097. In particular, Applicants have identified and isolated cDNA encoding a PRO1097 polypeptide, as disclosed in further detail in the Examples below. To Applicants present
25 knowledge, the DNA59841-1460 nucleotide sequence encodes a novel factor. Using BLAST and FastA sequence alignment computer programs, some sequence identity with proteins designated as "CELK05G3_3", "CRU26344_1", "SPBC16C6_8", "P_W13844" and "AF013403" was revealed.

105. Full-length PRO1153 Polypeptides

30 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1153. In particular, cDNA encoding a PRO1153 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

 Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1153 (shown in Figure 246 and SEQ ID NO:351) has certain amino acid sequence identity with
35 HPBR11-7 protein submitted to the EMBL Data Library June 1992. Accordingly, it is presently believed that PRO1153 disclosed in the present application may be related to HPBR11-7.

106. Full-length PRO1154 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1154. In particular, cDNA encoding a PRO1154 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native
5 sequence PRO1154 (shown in Figure 248 and SEQ ID NO:353) aligns with a KIAA0525 protein, designated ABO11097. PRO1154 has a novel N-terminus of 73 amino acids. Accordingly, PRO1154 is believed to be novel. PRO1154 also has significant sequence identity with aminopeptidase N, insulin-regulated membrane aminopeptidase, throtropin-releasing hormone degrading enzyme and placental leucine aminopeptidase. Therefore, PRO1154 is believed to be a novel aminopeptidase, or peptide which degrades peptides.

107. Full-length PRO1181 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1181. In particular, cDNA encoding a PRO1181 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

15 The DNA59847-1511 clone was isolated from a human prostate tissue library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the DNA59847-1511 clone does encode a secreted factor. As far as is known, the DNA59847-1511 sequence encodes a novel factor designated herein as PRO1181; using the WU-BLAST2 sequence alignment computer program, no significant sequence identities to any known proteins were revealed.

108. Full-length PRO1182 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1182. In particular, cDNA encoding a PRO1182 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

25 Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1182 (shown in Figure 252 and SEQ ID NO:357) has amino acid sequence identity with the conglutinin protein. Accordingly, it is presently believed that PRO1182 disclosed in the present application is a newly identified conglutinin homolog.

109. Full-length PRO1155 Polypeptides

30 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1155. In particular, cDNA encoding a PRO1155 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

35 Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1155 (shown in Figure 254 and SEQ ID NO:359) has certain amino acid sequence identity with neurokinin B. Accordingly, it is presently believed that PRO1155 disclosed in the present application is a newly identified member of the tachykinin family.

110. Full-length PRO1156 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1156. In particular, cDNA encoding a PRO1156 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

The DNA59853-1505 clone was isolated from an adult human heart library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the DNA59853-1505 clone may encode a secreted factor. As far as is known, the DNA59853-1505 sequence encodes a novel factor designated herein as PRO1156. However, using WU-BLAST2 sequence alignment computer programs, some sequence identity with known proteins were revealed.

111. Full-length PRO1098 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1098. In particular, Applicants have identified cDNA encoding a PRO1098 polypeptide, as disclosed in further detail in the Examples below. The PRO1098-encoding clone was isolated from a human lung tissue library. To Applicants present knowledge, the DNA59854-1459 nucleotide sequence encodes a novel factor; using BLAST and FastA sequence alignment computer programs, no significant sequence identities to any known proteins were revealed. Some sequence identity appeared with proteins such as the "Env" polyprotein and a methyltransferase.

112. Full-length PRO1127 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1127. In particular, cDNA encoding a PRO1127 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

The DNA60283-1484 clone encodes a secreted factor. As far as is known, the DNA60283-1484 sequence encodes a novel factor designated herein as PRO1127; using WU-BLAST2 sequence alignment computer programs, minimal sequence identities to any known proteins were revealed.

113. Full-length PRO1126 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1126. In particular, cDNA encoding a PRO1126 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1126 (shown in Figure 262 and SEQ ID NO:367) has certain amino acid sequence identity with the olfactomedin protein. Accordingly, it is presently believed that PRO1126 disclosed in the present application is a newly identified olfactomedin homolog and may possess activity typical of that protein.

114. Full-length PRO1125 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1125. In particular, cDNA encoding a PRO1125 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1125 (shown in Figure 264 and SEQ ID NO:369) has certain amino acid sequence identity with transcriptional repressor rco-1. Accordingly, it is presently believed that PRO1125 disclosed in the present application is a newly identified member of the WD superfamily.

115. Full-length PRO1186 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1186. In particular, cDNA encoding a PRO1186 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1186 (shown in Figure 266 and SEQ ID NO:371) has amino acid sequence identity with venom protein A from *Dendroaspis polylepis* venom. Accordingly, it is presently believed that PRO1186 disclosed in the present application is a newly identified member of venom protein A and may share a related mechanism.

116. Full-length PRO1198 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1198. In particular, cDNA encoding a PRO1198 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

As far as is known, the DNA60622-1525 sequence encodes a novel factor designated herein as PRO1198. However, using WU-BLAST2 sequence alignment computer programs, some sequence identity with known proteins was found.

117. Full-length PRO1158 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1158. In particular, cDNA encoding a PRO1158 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

The DNA60625-1507 clone was isolated from a human lung tumor tissue library. As far as is known, the DNA60625-1507 sequence encodes a novel factor designated herein as PRO1158. However, using WU-BLAST2 sequence alignment computer programs, some sequence identities with known proteins were shown.

118. Full-length PRO1159 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1159. In particular, cDNA encoding a PRO1159 polypeptide has

been identified and isolated, as disclosed in further detail in the Examples below.

The DNA60627-1508 clone was isolated from a human peripheral blood granulocyte tissue library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the DNA60627-1508 clone does encode a secreted factor. As far as is known, the DNA60627-1508 sequence encodes a novel factor designated herein as PRO1159; using the WU-BLAST2 sequence alignment computer program, no
5 sequence identities to any known proteins were revealed.

119. Full-length PRO1124 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1124. In particular, cDNA encoding a PRO1124 polypeptide has
10 been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1124 (shown in Figure 274 and SEQ ID NO:377) has amino acid sequence identity with an epithelial chloride channel protein from bos taurus. PRO1124 also has sequence identity with ECAM-1. Accordingly, it is presently believed that PRO1124 disclosed in the present application is a newly identified cell
15 membrane protein involved in communication of cells either through ion channels or cell adhesion molecules.

120. Full-length PRO1287 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1287. In particular, cDNA encoding a PRO1287 polypeptide has
20 been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1287 (shown in Figure 276 and SEQ ID NO:381) has amino acid sequence identity with the radical fringe protein from Gallus gallus (GGU82088_1). Accordingly, it is presently believed that PRO1287 disclosed in the present application is a newly identified fringe protein homolog and may possess activity typical
25 of the fringe protein.

121. Full-length PRO1312 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1312. In particular, cDNA encoding a PRO1312 polypeptide has
30 been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST2 sequence alignment computer programs, some sequence identities with known proteins were revealed, but were determined not to be significant. Therefore, as far as is known, the DNA61873-1574 sequence encodes a novel transmembrane protein designated herein as PRO1312.

122. Full-length PRO1192 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1192. In particular, cDNA encoding a PRO1192 polypeptide has
35

been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1192 (shown in Figure 280 and SEQ ID NO:389) has amino acid sequence identity with trout PO-like glycoprotein (GEN12838 IP1). Accordingly, it is presently believed that PRO1192 disclosed in the present application is a newly identified member of the myelin PO glycoprotein family.

5

123. Full-length PRO1160 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1160. In particular, cDNA encoding a PRO1160 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

10 The DNA62872-1509 clone was isolated from a human breast tissue library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the DNA62872-1509 clone does encode a secreted factor. As far as is known, the DNA62872-1509 sequence encodes a novel factor designated herein as PRO1160; using the WU-BLAST2 sequence alignment computer program, no significant sequence identities to any known proteins were revealed.

15

124. Full-length PRO1187 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1187. In particular, cDNA encoding a PRO1187 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

20 As far as is known, the DNA62876-1517 sequence encodes a novel factor designated herein as PRO1187; using WU-BLAST2 sequence alignment computer programs, no significant sequence identities to any known proteins were revealed.

125. Full-length PRO1185 Polypeptides

25 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1185. In particular, cDNA encoding a PRO1185 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

As far as is known, the DNA62881-1515 clone encodes a novel factor designated herein as PRO1185; using WU-BLAST2 sequence alignment computer programs, no significant sequence identities to any known proteins were revealed.

30

126. Full-length PRO1345 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1345. In particular, cDNA encoding a PRO1345 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

35

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1345 (shown in Figure 288 and SEQ ID NO:403) has amino acid sequence identity with

the C-type lectin homolog precursor protein of bos taurus (BTU22298_1). Accordingly, it is presently believed that PRO1345 disclosed in the present application is a newly identified member of the C-type lectin protein family and may possess activity typical of that family or of the tetranectin protein in particular.

127. Full-length PRO1245 Polypeptides

5 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1245. In particular, cDNA encoding a PRO1245 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

The DNA64884-1527 clone was identified using methods that selects for nucleotide sequences encoding secreted proteins. As far as is known, the DNA64884-1527 sequence encodes a novel secreted factor designated 10 herein as PRO1245. Using WU-BLAST2 sequence alignment computer programs, some sequence identities to known proteins were revealed; however, it was determined that they were not significant.

128. Full-length PRO1358 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides 15 referred to in the present application as PRO1358. In particular, cDNA encoding a PRO1358 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1358 (shown in Figure 292 and SEQ ID NO:410) has amino acid sequence identity with RASP-1. Accordingly, it is presently believed that PRO1358 disclosed in the present application is a newly identified 20 member of the serpin family of serine protease inhibitors and may possess serine protease inhibition activity, protein catabolism inhibitory activity and/or be associated with regeneration of tissue.

129. Full-length PRO1195 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides 25 referred to in the present application as PRO1195. In particular, cDNA encoding a PRO1195 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1195 (shown in Figure 294 and SEQ ID NO:412) has amino acid sequence identity with MMU28486_1, termed a proline rich acidic protein from Mus musculus, locus MMU28486, Accession: 30 U28486, database GBTRANS, submitted 06-JUN-1995 by John W. Kasik. Accordingly, it is presently believed that PRO1195 disclosed in the present application is a newly identified member of this protein family.

130. Full-length PRO1270 Polypeptides

35 The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1270. In particular, cDNA encoding a PRO1270 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1270 (shown in Figure 296 and SEQ ID NO:414) has amino acid sequence identity with the lectin protein (XLU86699_1) of *Xenopus laevis*. Accordingly, it is presently believed that PRO1270 disclosed in the present application is a newly identified member of the lectin protein family and may possess activity typical of that family.

5

131. Full-length PRO1271 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1271. In particular, cDNA encoding a PRO1271 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

10

As far as is known, the DNA66309-1538 sequence encodes a novel factor designated herein as PRO1271; using WU-BLAST2 sequence alignment computer programs, no significant sequence identities to any known proteins were revealed (results further described in the examples below).

132. Full-length PRO1375 Polypeptides

15

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1375. In particular, cDNA encoding a PRO1375 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1375 (shown in Figure 300 and SEQ ID NO:418) has amino acid sequence identity PUT2.

20

Accordingly, it is presently believed that PRO1375 disclosed in the present application has at least one related mechanism of PUT2.

133. Full-length PRO1385 Polypeptides

25

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1385. In particular, cDNA encoding a PRO1385 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

The DNA68869-1610 clone was isolated from a human tissue library using a trapping technique which selects for nucleotide sequences encoding secreted proteins. Thus, the DNA68869-1610 clone does encode a secreted factor. As far as is known, the DNA68869-1610 sequence encodes a novel factor designated herein as PRO1385; using the WU-BLAST2 sequence alignment computer program, no significant sequence identities to any known proteins were revealed.

30

134. Full-length PRO1387 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1387. In particular, cDNA encoding a PRO1387 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

35

Using the WU-BLAST2 sequence alignment computer program, it has been found that a full-length native sequence PRO1387 (shown in Figure 304 and SEQ ID NO:422) has amino acid sequence identity with the myelin p0 protein precursor (MYP0_HETFR). Accordingly, it is presently believed that PRO1387 disclosed in the present application is a newly identified member of the myelin protein family and may possess activity typical of that family.

5

135. Full-length PRO1384 Polypeptides

The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO1384. In particular, cDNA encoding a PRO1384 polypeptide has been identified and isolated, as disclosed in further detail in the Examples below.

10

Using WU-BLAST2 sequence alignment computer programs, it has been found that a full-length native sequence PRO1384 (shown in Figure 306 and SEQ ID NO:424) has amino acid sequence identity with NKG2-D (AF054819_1; Dayhoff database, version 35.45 SwissProt 35). Accordingly, it is presently believed that PRO1384 disclosed in the present application is a newly identified member of the NKG2 family and may possess MHC activation/inactivation activities typical of the NKG2 family.

15

B. PRO Variants

In addition to the full-length native sequence PRO polypeptides described herein, it is contemplated that PRO variants can be prepared. PRO variants can be prepared by introducing appropriate nucleotide changes into the PRO DNA, and/or by synthesis of the desired PRO polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

20

Variations in the native full-length sequence PRO or in various domains of the PRO described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO that results in a change in the amino acid sequence of the PRO as compared with the native sequence PRO. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

25

30

PRO polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native protein.

Certain fragments lack amino acid residues that are not essential for a desired biological activity of the PRO polypeptide.

PRO fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating PRO fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, PRO polypeptide fragments share at least one biological and/or immunological activity with the native PRO polypeptide disclosed herein.

In particular embodiments, conservative substitutions of interest are shown in Table 1 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 1, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 1

	<u>Original Residue</u>	<u>Exemplary Substitutions</u>	<u>Preferred Substitutions</u>
5	Ala (A)	val; leu; ile	val
	Arg (R)	lys; gln; asn	lys
	Asn (N)	gln; his; lys; arg	gln
	Asp (D)	glu	glu
	Cys (C)	ser	ser
10	Gln (Q)	asn	asn
	Glu (E)	asp	asp
	Gly (G)	pro; ala	ala
	His (H)	asn; gln; lys; arg	arg
	Ile (I)	leu; val; met; ala; phe;	
15		norleucine	leu
	Leu (L)	norleucine; ile; val;	
		met; ala; phe	ile
	Lys (K)	arg; gln; asn	arg
	Met (M)	leu; phe; ile	leu
20	Phe (F)	leu; val; ile; ala; tyr	leu
	Pro (P)	ala	ala
	Ser (S)	thr	thr
	Thr (T)	ser	ser
	Trp (W)	tyr; phe	tyr
25	Tyr (Y)	trp; phe; thr; ser	phe
	Val (V)	ile; leu; met; phe;	
		ala; norleucine	leu

Substantial modifications in function or immunological identity of the PRO polypeptide are accomplished

30 by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

(1) hydrophobic: norleucine, met, ala, val, leu, ile;

35 (2) neutral hydrophilic: cys, ser, thr;

(3) acidic: asp, glu;

(4) basic: asn, gln, his, lys, arg;

(5) residues that influence chain orientation: gly, pro; and

(6) aromatic: trp, tyr, phe.

40 Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO variant

DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244: 1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

10 C. Modifications of PRO

Covalent modifications of PRO are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the PRO. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

Other modifications include deamidation of glutaminyl and asparaginyl residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence PRO. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

Addition of glycosylation sites to the PRO polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO (for O-linked glycosylation sites). The PRO amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO comprises linking the PRO polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

The PRO of the present invention may also be modified in a way to form a chimeric molecule comprising PRO fused to another, heterologous polypeptide or amino acid sequence.

In one embodiment, such a chimeric molecule comprises a fusion of the PRO with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO. The presence of such epitope-tagged forms of the PRO can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an α -tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a PRO polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27, 1995.

D. Preparation of PRO

The description below relates primarily to production of PRO by culturing cells transformed or transfected with a vector containing PRO nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO. For instance, the PRO sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the PRO may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO.

1. Isolation of DNA Encoding PRO

DNA encoding PRO may be obtained from a cDNA library prepared from tissue believed to possess the PRO mRNA and to express it at a detectable level. Accordingly, human PRO DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

Libraries can be screened with probes (such as antibodies to the PRO or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ³²P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example, CaCl_2 , CaPO_4 , liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA*; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kar'*; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7*

ilvG kan^r; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer et al., Bio/Technology, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., J. Bacteriol., 737 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickerhamii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilum* (ATCC 36,906; Van den Berg et al., Bio/Technology, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus*; *yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., J. Basic Microbiol., 28:265-278 [1988]); *Candida*; *Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case et al., Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn et al., Gene, 26:205-221 [1983]; Yelton et al., Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 [1985]). Methylotrophic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylotrophs, 269 (1982).

Suitable host cells for the expression of glycosylated PRO are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells.

More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen Virol., 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

3. Selection and Use of a Replicable Vector

The nucleic acid (e.g., cDNA or genomic DNA) encoding PRO may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally

include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

5 The PRO may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g.,
10 the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral
15 secretory leaders.

Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for
20 cloning vectors in mammalian cells.

Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

25 An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature,
30 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp1* gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

Expression and cloning vectors usually contain a promoter operably linked to the PRO-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well
35 known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (*trp*) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid

promoters such as the tac promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding PRO.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding the PRO by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO DNA and encoding a specific antibody epitope.

5. Purification of Polypeptide

Forms of PRO may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO produced.

E. Uses for PRO

Nucleotide sequences (or their complement) encoding PRO have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO nucleic acid will also be useful for the preparation of PRO polypeptides by the recombinant techniques described herein.

The full-length native sequence PRO gene, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO cDNA or to isolate still other cDNAs (for instance, those encoding naturally-occurring variants of PRO or PRO from other species) which have a desired sequence identity to the native PRO sequence disclosed herein. Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from at least partially novel regions of the full length native nucleotide sequence wherein those regions may be determined without undue experimentation or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO. By way of example, a screening method will comprise isolating the coding region of the PRO gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S , or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

Any EST sequences disclosed in the present application may similarly be employed as probes, using the methods disclosed herein.

Other useful fragments of the PRO nucleic acids include antisense or sense oligonucleotides comprising a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target PRO mRNA (sense) or PRO DNA (antisense) sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of PRO DNA. Such a fragment generally comprises at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (Cancer Res. 48:2659, 1988) and van der Krol et al. (BioTechniques 6:958, 1988).

Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of duplexes that block transcription or translation of the target sequence by one of several means, including enhanced degradation of the duplexes, premature termination of transcription or translation, or by other means. The antisense oligonucleotides thus may be used to block expression of PRO proteins. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages, such as those described in WO 91/06629) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable *in vivo* (i.e., capable of resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences.

Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increases affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

Antisense or sense oligonucleotides may be introduced into a cell containing the target nucleic acid sequence by any gene transfer method, including, for example, CaPO_4 -mediated DNA transfection, electroporation, or by using gene transfer vectors such as Epstein-Barr virus. In a preferred procedure, an antisense or sense oligonucleotide is inserted into a suitable retroviral vector. A cell containing the target nucleic acid sequence is contacted with the recombinant retroviral vector, either *in vivo* or *ex vivo*. Suitable retroviral vectors include, but are not limited to, those derived from the murine retrovirus M-MuLV, N2 (a retrovirus derived from M-MuLV), or the double copy vectors designated DCT5A, DCT5B and DCT5C (see WO 90/13641).

Sense or antisense oligonucleotides also may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell.

Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. The sense or antisense oligonucleotide-lipid complex is preferably dissociated within the cell by an endogenous lipase.

The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO coding sequences.

Nucleotide sequences encoding a PRO can also be used to construct hybridization probes for mapping the gene which encodes that PRO and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for PRO encode a protein which binds to another protein (example, where the PRO is a receptor), the PRO can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO or a receptor for PRO. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode PRO or its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene,

which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of PRO can be used to construct a PRO "knock out" animal which has a defective or altered gene encoding PRO as a result of homologous recombination between the endogenous gene encoding PRO and altered genomic DNA encoding PRO introduced into an embryonic stem cell of the animal. For example, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques. A portion of the genomic DNA encoding PRO can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, *Cell*, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., *Cell*, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO polypeptide.

Nucleic acid encoding the PRO polypeptides may also be used in gene therapy. In gene therapy applications, genes are introduced into cells in order to achieve *in vivo* synthesis of a therapeutically effective genetic product, for example for replacement of a defective gene. "Gene therapy" includes both conventional gene therapy where a lasting effect is achieved by a single treatment, and the administration of gene therapeutic agents, which involves the one time or repeated administration of a therapeutically effective DNA or mRNA.

Antisense RNAs and DNAs can be used as therapeutic agents for blocking the expression of certain genes *in vivo*. It has already been shown that short antisense oligonucleotides can be imported into cells where they act as inhibitors, despite their low intracellular concentrations caused by their restricted uptake by the cell membrane. (Zamecnik *et al.*, Proc. Natl. Acad. Sci. USA 83:4143-4146 [1986]). The oligonucleotides can be modified to enhance their uptake, e.g. by substituting their negatively charged phosphodiester groups by uncharged groups.

There are a variety of techniques available for introducing nucleic acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or *in vivo* in the cells of the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, the calcium phosphate precipitation method, etc. The currently preferred *in vivo* gene transfer techniques include transfection with viral (typically retroviral) vectors and viral coat protein-liposome mediated transfection (Dzau *et al.*, Trends in Biotechnology 11, 205-210 [1993]). In some situations it is desirable to provide the nucleic acid source with an agent that targets the target cells, such as an antibody specific for a cell surface membrane protein or the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins which bind to a cell surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, e.g. capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half-life. The technique of receptor-mediated endocytosis is described, for example, by Wu *et al.*, J. Biol. Chem. 262, 4429-4432 (1987); and Wagner *et al.*, Proc. Natl. Acad. Sci. USA 87, 3410-3414 (1990). For review of gene marking and gene therapy protocols see Anderson *et al.*, Science 256, 808-813 (1992).

The PRO polypeptides described herein may also be employed as molecular weight markers for protein electrophoresis purposes.

The nucleic acid molecules encoding the PRO polypeptides or fragments thereof described herein are useful for chromosome identification. In this regard, there exists an ongoing need to identify new chromosome markers, since relatively few chromosome marking reagents, based upon actual sequence data are presently available. Each PRO nucleic acid molecule of the present invention can be used as a chromosome marker.

The PRO polypeptides and nucleic acid molecules of the present invention may also be used for tissue typing, wherein the PRO polypeptides of the present invention may be differentially expressed in one tissue as compared to another. PRO nucleic acid molecules will find use for generating probes for PCR, Northern analysis, Southern analysis and Western analysis.

The PRO polypeptides described herein may also be employed as therapeutic agents. The PRO polypeptides of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the PRO product hereof is combined in admixture with a pharmaceutically acceptable carrier vehicle. Therapeutic formulations are prepared for storage by mixing the active ingredient having the desired degree of purity with optional physiologically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients or stabilizers are nontoxic to recipients at the

dosages and concentrations employed, and include buffers such as phosphate, citrate and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone, amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, PLURONICS™ or PEG.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes, prior to or following lyophilization and reconstitution.

Therapeutic compositions herein generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

The route of administration is in accord with known methods, e.g. injection or infusion by intravenous, intraperitoneal, intracerebral, intramuscular, intraocular, intraarterial or intralesional routes, topical administration, or by sustained release systems.

Dosages and desired drug concentrations of pharmaceutical compositions of the present invention may vary depending on the particular use envisioned. The determination of the appropriate dosage or route of administration is well within the skill of an ordinary physician. Animal experiments provide reliable guidance for the determination of effective doses for human therapy. Interspecies scaling of effective doses can be performed following the principles laid down by Mordenti, J. and Chappell, W. "The use of interspecies scaling in toxicokinetics" In *Toxicokinetics and New Drug Development*, Yacobi et al., Eds., Pergamon Press, New York 1989, pp. 42-96.

When *in vivo* administration of a PRO polypeptide or agonist or antagonist thereof is employed, normal dosage amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, preferably about 1 µg/kg/day to 10 mg/kg/day, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting one organ or tissue, for example, may necessitate delivery in a manner different from that to another organ or tissue.

Where sustained-release administration of a PRO polypeptide is desired in a formulation with release characteristics suitable for the treatment of any disease or disorder requiring administration of the PRO polypeptide, microencapsulation of the PRO polypeptide is contemplated. Microencapsulation of recombinant proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon- (rhIFN-), interleukin-2, and MN rgp120. Johnson et al., Nat. Med., 2:795-799 (1996); Yasuda, Biomed. Ther., 27:1221-1223 (1993); Hora et al., Bio/Technology, 8:755-758 (1990); Cleland, "Design and Production of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in Vaccine Design: The Subunit and Adjuvant Approach, Powell and Newman, eds, (Plenum Press: New York, 1995), pp. 439-462; WO 97/03692, WO 96/40072, WO 96/07399; and U.S. Pat. No. 5,654,010.

The sustained-release formulations of these proteins were developed using poly-lactic-co-glycolic acid (PLGA) polymer due to its biocompatibility and wide range of biodegradable properties. The degradation products of PLGA, lactic and glycolic acids, can be cleared quickly within the human body. Moreover, the degradability of this polymer can be adjusted from months to years depending on its molecular weight and composition. Lewis, "Controlled release of bioactive agents from lactide/glycolide polymer," in: M. Chasin and R. Langer (Eds.), Biodegradable Polymers as Drug Delivery Systems (Marcel Dekker: New York, 1990), pp. 1-41.

This invention encompasses methods of screening compounds to identify those that mimic the PRO polypeptide (agonists) or prevent the effect of the PRO polypeptide (antagonists). Screening assays for antagonist drug candidates are designed to identify compounds that bind or complex with the PRO polypeptides encoded by the genes identified herein, or otherwise interfere with the interaction of the encoded polypeptides with other cellular proteins. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates.

The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays, and cell-based assays, which are well characterized in the art.

All assays for antagonists are common in that they call for contacting the drug candidate with a PRO polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

In binding assays, the interaction is binding and the complex formed can be isolated or detected in the reaction mixture. In a particular embodiment, the PRO polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, e.g., on a microtiter plate, by covalent or non-covalent attachments. Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the PRO polypeptide and drying. Alternatively, an immobilized antibody, e.g., a monoclonal antibody, specific for the PRO polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, e.g., the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, e.g., by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

If the candidate compound interacts with but does not bind to a particular PRO polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, e.g., cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers (Fields and Song, Nature (London), 340:245-246 (1989); Chien et al., Proc. Natl. Acad. Sci. USA, 88:9578-9582 (1991)) as disclosed by Chevray and Nathans, Proc. Natl. Acad. Sci. USA, 89: 5789-5793 (1991). Many

transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system") takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-*lacZ* reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for β -galactosidase. A complete kit (MATCHMAKER™) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

Compounds that interfere with the interaction of a gene encoding a PRO polypeptide identified herein and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation) between the test compound and the intra- or extracellular component present in the mixture is monitored as described hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

To assay for antagonists, the PRO polypeptide may be added to a cell along with the compound to be screened for a particular activity and the ability of the compound to inhibit the activity of interest in the presence of the PRO polypeptide indicates that the compound is an antagonist to the PRO polypeptide. Alternatively, antagonists may be detected by combining the PRO polypeptide and a potential antagonist with membrane-bound PRO polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The PRO polypeptide can be labeled, such as by radioactivity, such that the number of PRO polypeptide molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Coligan et al., Current Protocols in Immun., 1(2): Chapter 5 (1991). Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the PRO polypeptide and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the PRO polypeptide. Transfected cells that are grown on glass slides are exposed to labeled PRO polypeptide. The PRO polypeptide can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an interactive sub-pooling and re-screening process, eventually yielding a

single clone that encodes the putative receptor.

As an alternative approach for receptor identification, labeled PRO polypeptide can be photoaffinity-linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE and exposed to X-ray film. The labeled complex containing the receptor can be excised, resolved into peptide fragments, and subjected to protein micro-sequencing. The amino acid sequence obtained
5 from micro-sequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

In another assay for antagonists, mammalian cells or a membrane preparation expressing the receptor would be incubated with labeled PRO polypeptide in the presence of the candidate compound. The ability of the compound to enhance or block this interaction could then be measured.

10 More specific examples of potential antagonists include an oligonucleotide that binds to the fusions of immunoglobulin with PRO polypeptide, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. Alternatively, a potential antagonist may be a closely related protein, for example, a mutated form of the PRO
15 polypeptide that recognizes the receptor but imparts no effect, thereby competitively inhibiting the action of the PRO polypeptide.

Another potential PRO polypeptide antagonist is an antisense RNA or DNA construct prepared using antisense technology, where, e.g., an antisense RNA or DNA molecule acts to block directly the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense technology can be used
20 to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes the mature PRO polypeptides herein, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res.,
25 6:3073 (1979); Cooney et al., Science, 241: 456 (1988); Dervan et al., Science, 251:1360 (1991)), thereby preventing transcription and the production of the PRO polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the PRO polypeptide (antisense - Okano, Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, FL, 1988). The oligonucleotides described above can also be delivered
30 to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the PRO polypeptide. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation-initiationsite, e.g., between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

Potential antagonists include small molecules that bind to the active site, the receptor binding site, or growth factor or other relevant binding site of the PRO polypeptide, thereby blocking the normal biological
35 activity of the PRO polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details see, e.g., Rossi, Current Biology, 4:469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

5 Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation via Hoogsteen base-pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details see, e.g., PCT publication No. WO 97/33551, *supra*.

10 These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

PRO189 can be used in assays with W01A6.1 of *C. Elegans*, phosphodiesterases, transporters and proteins which bind to fatty acids, to determine the relative activities of PRO189 against these proteins. The results can be applied accordingly.

15

F. Anti-PRO Antibodies

The present invention further provides anti-PRO antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

20

1. Polyclonal Antibodies

The anti-PRO antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

25

30 The immunization protocol may be selected by one skilled in the art without undue experimentation.

2. Monoclonal Antibodies

The anti-PRO antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975).

35 In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.

The immunizing agent will typically include the PRO polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103].

- 5 Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will
- 10 include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

- Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk
- 15 Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

- The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of
- 20 monoclonal antibodies directed against PRO. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

- 25 After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, supra]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

- The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture
- 30 medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

- The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding
- 35 specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells,

or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., *supra*] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

3. Human and Humanized Antibodies

The anti-PRO antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.*, 2:593-596 (1992)].

Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers

[Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven et al., Science, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, p. 77 (1985) and Boerner et al., J. Immunol., 147(1):86-95 (1991)). Similarly, human antibodies can be made by introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks et al., Bio/Technology 10, 779-783 (1992); Lonberg et al., Nature 368 856-859 (1994); Morrison, Nature 368, 812-13 (1994); Fishwild et al., Nature Biotechnology 14, 845-51 (1996); Neuberger, Nature Biotechnology 14, 826 (1996); Lonberg and Huszar, Intern. Rev. Immunol. 13 65-93 (1995).

4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the PRO, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, Nature, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in

at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')₂ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Fab' fragments may be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various technique for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with

the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber *et al.*, J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt *et al.*, J. Immunol. 147:60 (1991).

- 5 Exemplary bispecific antibodies may bind to two different epitopes on a given PRO polypeptide herein. Alternatively, an anti-PRO polypeptide arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular PRO polypeptide. Bispecific antibodies may also be used to localize
10 cytotoxic agents to cells which express a particular PRO polypeptide. These antibodies possess a PRO-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the PRO polypeptide and further binds tissue factor (TF).

15 5. Heteroconjugate Antibodies

- Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro*
20 using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

25 6. Effector Function Engineering

- It may be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) may be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased
30 complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron *et al.*, J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff *et al.* Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson *et al.*, Anti-Cancer Drug Design, 3: 219-230 (1989).
35

7. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above.

- 5 Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolacca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are
10 available for the production of radioconjugated antibodies. Examples include ^{212}Bi , ^{131}I , ^{131}In , ^{90}Y , and ^{186}Re .

- Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-
15 diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

- 20 In another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is conjugated to a cytotoxic agent (e.g., a radionucleotide).

25 8. Immunoliposomes

- The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein *et al.*, Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent No.
30 5,013,556.

- Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin
35 *et al.*, J. Biol. Chem., 257: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. See Gabizon *et al.*, J. National Cancer Inst., 81(19): 1484 (1989).

9. Pharmaceutical Compositions of Antibodies

Antibodies specifically binding a PRO polypeptide identified herein, as well as other molecules identified by the screening assays disclosed hereinbefore, can be administered for the treatment of various disorders in the form of pharmaceutical compositions.

If the PRO polypeptide is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, lipofections or liposomes can also be used to deliver the antibody, or an antibody fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, *e.g.*, Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993). The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles, and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, *supra*.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOTTM (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated antibodies remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulfhydryl residues, lyophilizing from acidic solutions, controlling moisture content, using appropriate additives, and developing specific polymer matrix compositions.

G. Uses for anti-PRO Antibodies

The anti-PRO antibodies of the invention have various utilities. For example, anti-PRO antibodies may be used in diagnostic assays for PRO, *e.g.*, detecting its expression in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases

5 [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as ^3H , ^{14}C , ^{32}P , ^{35}S , or ^{125}I , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or
10 horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

Anti-PRO antibodies also are useful for the affinity purification of PRO from recombinant cell culture
15 or natural sources. In this process, the antibodies against PRO are immobilized on a suitable support, such a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO from the
20 antibody.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas,
30 VA.

EXAMPLE 1: Extracellular Domain Homology Screening to Identify Novel Polypeptides and cDNA Encoding Therefor

The extracellular domain (ECD) sequences (including the secretion signal sequence, if any) from about
35 950 known secreted proteins from the Swiss-Prot public database were used to search EST databases. The EST databases included public databases (*e.g.*, Dayhoff, GenBank), and proprietary databases (*e.g.* LIFESEQTM, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program WU-BLAST-2

(Altschul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. Those comparisons with a Blast score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, WA).

Using this extracellular domain homology screen, consensus DNA sequences were assembled relative to the other identified EST sequences using phrap. In addition, the consensus DNA sequences obtained were often (but not always) extended using repeated cycles of WU-BLAST-2 and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above.

Based upon the consensus sequences obtained as described above, oligonucleotides were then synthesized and used to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for a PRO polypeptide. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

EXAMPLE 2: Isolation of cDNA clones by Amylase Screening

1. Preparation of oligo dT primed cDNA library

mRNA was isolated from a human tissue of interest using reagents and protocols from Invitrogen, San Diego, CA (Fast Track 2). This RNA was used to generate an oligo dT primed cDNA library in the vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). In this procedure, the double stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linked cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

2. Preparation of random primed cDNA library

A secondary cDNA library was generated in order to preferentially represent the 5' ends of the primary cDNA clones. Sp6 RNA was generated from the primary library (described above), and this RNA was used to generate a random primed cDNA library in the vector pSST-AMY.0 using reagents and protocols from Life Technologies (Super Script Plasmid System, referenced above). In this procedure the double stranded cDNA

was sized to 500-1000 bp, linked with blunt to NotI adaptors, cleaved with SfiI, and cloned into SfiI/NotI cleaved vector. pSST-AMY.0 is a cloning vector that has a yeast alcohol dehydrogenase promoter preceding the cDNA cloning sites and the mouse amylase sequence (the mature sequence without the secretion signal) followed by the yeast alcohol dehydrogenase terminator, after the cloning sites. Thus, cDNAs cloned into this vector that are fused in frame with amylase sequence will lead to the secretion of amylase from appropriately transfected yeast colonies.

3. Transformation and Detection

DNA from the library described in paragraph 2 above was chilled on ice to which was added electrocompetent DH10B bacteria (Life Technologies, 20 ml). The bacteria and vector mixture was then electroporated as recommended by the manufacturer. Subsequently, SOC media (Life Technologies, 1 ml) was added and the mixture was incubated at 37°C for 30 minutes. The transformants were then plated onto 20 standard 150 mm LB plates containing ampicillin and incubated for 16 hours (37°C). Positive colonies were scraped off the plates and the DNA was isolated from the bacterial pellet using standard protocols, e.g. CsCl-gradient. The purified DNA was then carried on to the yeast protocols below.

The yeast methods were divided into three categories: (1) Transformation of yeast with the plasmid/cDNA combined vector; (2) Detection and isolation of yeast clones secreting amylase; and (3) PCR amplification of the insert directly from the yeast colony and purification of the DNA for sequencing and further analysis.

The yeast strain used was HD56-5A (ATCC-90785). This strain has the following genotype: MAT alpha, ura3-52, leu2-3, leu2-112, his3-11, his3-15, MAL⁺, SUC⁺, GAL⁺. Preferably, yeast mutants can be employed that have deficient post-translational pathways. Such mutants may have translocation deficient alleles in *sec71*, *sec72*, *sec62*, with truncated *sec71* being most preferred. Alternatively, antagonists (including antisense nucleotides and/or ligands) which interfere with the normal operation of these genes, other proteins implicated in this post translation pathway (e.g., SEC61p, SEC72p, SEC62p, SEC63p, TDJ1p or SSA1p-4p) or the complex formation of these proteins may also be preferably employed in combination with the amylase-expressing yeast.

Transformation was performed based on the protocol outlined by Gietz et al., Nucl. Acid. Res., 20:1425 (1992). Transformed cells were then inoculated from agar into YEPD complex media broth (100 ml) and grown overnight at 30°C. The YEPD broth was prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 207 (1994). The overnight culture was then diluted to about 2×10^6 cells/ml (approx. OD₆₀₀=0.1) into fresh YEPD broth (500 ml) and regrown to 1×10^7 cells/ml (approx. OD₆₀₀=0.4-0.5).

The cells were then harvested and prepared for transformation by transfer into GS3 rotor bottles in a Sorval GS3 rotor at 5,000 rpm for 5 minutes, the supernatant discarded, and then resuspended into sterile water, and centrifuged again in 50 ml falcon tubes at 3,500 rpm in a Beckman GS-6KR centrifuge. The supernatant was discarded and the cells were subsequently washed with LiAc/TE (10 ml, 10 mM Tris-HCl, 1 mM EDTA pH 7.5, 100 mM Li₂OOCCH₃), and resuspended into LiAc/TE (2.5 ml).

Transformation took place by mixing the prepared cells (100 μ l) with freshly denatured single stranded salmon testes DNA (Lofstrand Labs, Gaithersburg, MD) and transforming DNA (1 μ g, vol. < 10 μ l) in microfuge tubes. The mixture was mixed briefly by vortexing, then 40% PEG/TE (600 μ l, 40% polyethylene glycol-4000, 10 mM Tris-HCl, 1 mM EDTA, 100 mM Li₂OOCCH₃, pH 7.5) was added. This mixture was gently mixed and incubated at 30°C while agitating for 30 minutes. The cells were then heat shocked at 42°C for 15 minutes, and the reaction vessel centrifuged in a microfuge at 12,000 rpm for 5-10 seconds, decanted and resuspended into TE (500 μ l, 10 mM Tris-HCl, 1 mM EDTA pH 7.5) followed by recentrifugation. The cells were then diluted into TE (1 ml) and aliquots (200 μ l) were spread onto the selective media previously prepared in 150 mm growth plates (VWR).

Alternatively, instead of multiple small reactions, the transformation was performed using a single, large scale reaction, wherein reagent amounts were scaled up accordingly.

The selective media used was a synthetic complete dextrose agar lacking uracil (SCD-Ura) prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 208-210 (1994). Transformants were grown at 30°C for 2-3 days.

The detection of colonies secreting amylase was performed by including red starch in the selective growth media. Starch was coupled to the red dye (Reactive Red-120, Sigma) as per the procedure described by Biely et al., Anal. Biochem., 172:176-179 (1988). The coupled starch was incorporated into the SCD-Ura agar plates at a final concentration of 0.15% (w/v), and was buffered with potassium phosphate to a pH of 7.0 (50-100 mM final concentration).

The positive colonies were picked and streaked across fresh selective media (onto 150 mm plates) in order to obtain well isolated and identifiable single colonies. Well isolated single colonies positive for amylase secretion were detected by direct incorporation of red starch into buffered SCD-Ura agar. Positive colonies were determined by their ability to break down starch resulting in a clear halo around the positive colony visualized directly.

4. Isolation of DNA by PCR Amplification

When a positive colony was isolated, a portion of it was picked by a toothpick and diluted into sterile water (30 μ l) in a 96 well plate. At this time, the positive colonies were either frozen and stored for subsequent analysis or immediately amplified. An aliquot of cells (5 μ l) was used as a template for the PCR reaction in a 25 μ l volume containing: 0.5 μ l KlenTaq (Clontech, Palo Alto, CA); 4.0 μ l 10 mM dNTP's (Perkin Elmer-Cetus); 2.5 μ l KlenTaq buffer (Clontech); 0.25 μ l forward oligo 1; 0.25 μ l reverse oligo 2; 12.5 μ l distilled water. The sequence of the forward oligonucleotide 1 was:

5'-TGTAACACGACGGCCAGTTAAATAGACCTGCAATTATTAATCT-3' (SEQ ID NO:3)

The sequence of reverse oligonucleotide 2 was:

5'-CAGGAAACAGCTATGACCACCTGCACACCTGCAAATCCATT-3' (SEQ ID NO:4)

PCR was then performed as follows:

- a. Denature 92°C, 5 minutes
- b. 3 cycles of: Denature 92°C, 30 seconds

		Anneal	59°C, 30 seconds
		Extend	72°C, 60 seconds
5	c. 3 cycles of:	Denature	92°C, 30 seconds
		Anneal	57°C, 30 seconds
		Extend	72°C, 60 seconds
	d. 25 cycles of:	Denature	92°C, 30 seconds
		Anneal	55°C, 30 seconds
		Extend	72°C, 60 seconds
10	e.	Hold	4°C

The underlined regions of the oligonucleotides annealed to the ADH promoter region and the amylase region, respectively, and amplified a 307 bp region from vector pSST-AMY.0 when no insert was present. Typically, the first 18 nucleotides of the 5' end of these oligonucleotides contained annealing sites for the sequencing primers. Thus, the total product of the PCR reaction from an empty vector was 343 bp. However, signal sequence-fused cDNA resulted in considerably longer nucleotide sequences.

Following the PCR, an aliquot of the reaction (5 μ l) was examined by agarose gel electrophoresis in a 1% agarose gel using a Tris-Borate-EDTA (TBE) buffering system as described by Sambrook et al., supra. Clones resulting in a single strong PCR product larger than 400 bp were further analyzed by DNA sequencing after purification with a 96 Qiaquick PCR clean-up column (Qiagen Inc., Chatsworth, CA).

EXAMPLE 3: Isolation of cDNA Clones Using Signal Algorithm Analysis

Various polypeptide-encoding nucleic acid sequences were identified by applying a proprietary signal sequence finding algorithm developed by Genentech, Inc. (South San Francisco, CA) upon ESTs as well as clustered and assembled EST fragments from public (e.g., GenBank) and/or private (LIFESEQ®, Incyte Pharmaceuticals, Inc., Palo Alto, CA) databases. The signal sequence algorithm computes a secretion signal score based on the character of the DNA nucleotides surrounding the first and optionally the second methionine codon(s) (ATG) at the 5'-end of the sequence or sequence fragment under consideration. The nucleotides following the first ATG must code for at least 35 unambiguous amino acids without any stop codons. If the first ATG has the required amino acids, the second is not examined. If neither meets the requirement, the candidate sequence is not scored. In order to determine whether the EST sequence contains an authentic signal sequence, the DNA and corresponding amino acid sequences surrounding the ATG codon are scored using a set of seven sensors (evaluation parameters) known to be associated with secretion signals. Use of this algorithm resulted in the identification of numerous polypeptide-encoding nucleic acid sequences.

EXAMPLE 4: Isolation of cDNA clones Encoding Human PRO281

In order to obtain a cDNA clone encoding PRO281, methods described in Klein et al., Proc. Natl. Acad. Sci. USA 93:7108-7113 (1996) were employed with the following modifications. Yeast transformation was performed with limiting amounts of transforming DNA in order to reduce the number of multiple transformed yeast cells. Instead of plasmid isolation from the yeast followed by transformation of *E. coli* as

described in Klein et al., *supra*, PCR analysis was performed on single yeast colonies. PCR primers employed were bipartite in order to amplify the insert and a small portion of the invertase gene (allowing to determine that the insert was in frame with invertase) and to add on universal sequencing primer sites.

An invertase library was transformed into yeast and positives were selected on sucrose plates. Positive clones were re-tested and PCR products were sequenced. The sequence of one clone, PRO281, was determined to contain a signal peptide coding sequence. Oligonucleotide primers and probes were designed using the nucleotide sequence of PRO281. A full length plasmid library of cDNAs from human umbilical vein endothelium tissue was titered and approximately 100,000 cfu were plated in 192 pools of 500 cfu/pool into 96-well round bottom plates. The plates were sealed and pools were grown overnight at 37°C with shaking (200rpm). PCR was performed on the individual cultures using primers. Agarose gel electrophoresis was performed and positive wells were identified by visualization of a band of the expected size. Individual positive clones were obtained by colony lift followed by hybridization with ³²P-labeled oligonucleotide. These clones were characterized by PCR, restriction digest, and southern blot analyses.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 80-82, and a stop signal at nucleotide positions 1115-1117 (Figure 1, SEQ ID NO:1). The predicted polypeptide precursor is 345 amino acids long, has a calculated molecular weight of approximately 37,205 daltons and an estimated pI of approximately 10.15. Analysis of the full-length PRO281 sequence shown in Figure 2 (SEQ ID NO:2) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 14, multiple transmembrane domains from about amino acid position 83 to about amino acid position 105, from about amino acid position 126 to about amino acid position 146, from about amino acid position 158 to about amino acid position 177, from about amino acid position 197 to about amino acid position 216, from about amino acid position 218 to about amino acid position 238, from about amino acid position 245 to about amino acid position 265, and from about amino acid position 271 to about amino acid position 290 and an amino acid sequence block having homology to G-protein coupled receptor proteins from about amino acid 115 to about amino acid 155. Clone UNQ244 (DNA16422-1209) has been deposited with ATCC on June 2, 1998 and is assigned ATCC deposit no. 209929.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 2 (SEQ ID NO:2), evidenced significant homology between the PRO281 amino acid sequence and the following Dayhoff sequences: H64634, AF033095_1, B64815, YBHL_ECOLI, EMEQUTR_1, AF064763_3, S53708, A69253, AF035413_12 and S63281.

EXAMPLE 5: Isolation of cDNA clones Encoding Human PRO276

In order to obtain a cDNA clone encoding PRO276, methods described in Klein et al., *PNAS*, 93:7108-7113 (1996) were employed with the following modifications. Yeast transformation was performed with limiting amounts of transforming DNA in order to reduce the number of multiple transformed yeast cells. Instead of plasmid isolation from the yeast followed by transformation of *E. coli* as described in Klein et al., *supra*, PCR analysis was performed on single yeast colonies. PCR primers employed were bipartite in order to amplify the

insert and a small portion of the invertase gene (allowing to determine that the insert was in frame with invertase) and to add on universal sequencing primer sites.

An invertase library was transformed into yeast and positives were selected on sucrose plates. Positive clones were re-tested and PCR products were sequenced. The sequence of one clone, PRO276, was determined to contain a signal peptide coding sequence. Oligonucleotide primers and probes were designed using the nucleotide sequence of PRO276. A full length plasmid library of cDNAs from human fetal liver cells was titrated and approximately 100,000 cfu were plated in 192 pools of 500 cfu/pool into 96-well round bottom plates. The plates were sealed and pools were grown overnight at 37 C with shaking (200rpm). PCR was performed on the individual cultures using primers. Agarose gel electrophoresis was performed and positive wells were identified by visualization of a band of the expected size. Individual positive clones were obtained by colony lift followed by hybridization with ³²P-labeled oligonucleotide. These clones were characterized by PCR, restriction digest, and southern blot analyses.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 180-182 and a stop signal at nucleotide positions 933-935 (Figure 3; SEQ ID NO:5). The predicted polypeptide precursor is 251 amino acids long has a calculated molecular weight of approximately 28,801 daltons and an estimated pI of approximately 9.58. The transmembrane domains are approximately at amino acids 98-116 and 152-172 of the sequence shown in Figure 4 (SEQ ID NO:6). Clone DNA16435-1208 (UNQ243) has been deposited with the ATCC and is assigned ATCC deposit no. 209930.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 4 (SEQ ID NO:6), revealed some sequence identity between the PRO276 amino acid sequence and the following Dayhoff sequences: CEG25D7_2, ATT8O5_2, S69696, GRHR_RAT, NPCBAABCD_3, AB013149_1, P_R85942 and AP000006_5.

EXAMPLE 6: Isolation of cDNA clones Encoding Human PRO189

A clone designated herein as DNA14187 was isolated as described in Example 2 above from a human retina tissue library. The DNA14187 sequence is shown in Figure 7 (SEQ ID NO:9). Based on the DNA14187 sequence shown in Figure 7 (SEQ ID NO:9), oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO189. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TTGACCTATACAGAGATTCATC-3' (SEQ ID NO:10); and
reverse PCR primer 5'-CTAAGAACTTCCTCAGGATTTT-3' (SEQ ID NO:11).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA14187 sequence which had the following nucleotide sequence:

hybridization probe

5'-ATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGC-3' (SEQ ID NO:12).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO189 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human retina tissue (LIB94). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to Sall hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO189 and the derived protein sequence for PRO189.

The entire nucleotide sequence of DNA21624-1391 is shown in Figure 5 (SEQ ID NO:7). Clone DNA21624-1391 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 200-202 and ending at the stop codon at nucleotide positions 1301-1303 (Figure 5). The predicted polypeptide precursor is 367 amino acids long (Figure 6). The full-length PRO189 protein shown in Figure 6 has an estimated molecular weight of about 41,871 daltons and a pI of about 5.06. Clone DNA21624-1391 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analyzing the amino acid sequence of SEQ ID NO:8, the putative N-glycosylation sites are at about amino acids 224-227, 246-249 and 285-288. A domain for cytosolic fatty-acid binding proteins is at amino acids 78-107 of SEQ ID NO:8. The corresponding nucleotides can be routinely determined given the sequences provided herein.

Some sequence identity was found to W01A6.1 and F35D11.11, C. Elegans proteins, designated in a Dayhoff database as CEW01A6_10 and CELF35D11_11, respectively. Some sequence identity was also found to an antigen to malaria and to restin, designated in a Dayhoff database as P_R05766 and AF014012_1, respectively. Some sequence identity was also found to a microtubule binding protein and to myosin, designated in a Dayhoff database as AF041382_1 and S07537, respectively. There is also some sequence identity with 1-phosphatidylinositol-4, 5-bisphosphate, designated as PIP1_RAT.

EXAMPLE 7: Isolation of cDNA clones Encoding Human PRO190

A clone designated herein as DNA14232 was isolated as described in Example 2 above from a human fetal retina tissue library. The DNA14232 sequence is shown in Figure 10 (SEQ ID NO:15). Based on the DNA14232 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained

the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO190. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CTATACCTACTGTAGCTTCT-3' (SEQ ID NO:16); and

reverse PCR primer 5'-TCAGAGAATTCCTTCCAGGA-3' (SEQ ID NO:17).

- 10 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA14232 sequence which had the following nucleotide sequence:

hybridization probe

5'-ACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCACAACA-3' (SEQ ID NO:18).

- 15 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO190 gene using the probe oligonucleotide and one of the PCR primers.

- RNA for construction of the cDNA libraries was isolated from human retina tissue (LIB94). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

- 25 DNA sequencing of the clones isolated as described above gave sequences which include the full-length DNA sequence for PRO190 [herein designated as DNA23334-1392] (SEQ ID NO:13) and the derived protein sequence for PRO190.

- The entire nucleotide sequence of DNA23334-1392 is shown in Figure 8 (SEQ ID NO:13). Clone DNA23334-1392 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 193-195 and which ends at the stop codon at nucleotide positions 1465-1467 (Figure 8). The predicted polypeptide precursor is 424 amino acids long (Figure 9). The full-length PRO190 protein shown in Figure 9 has an estimated molecular weight of about 48,500 daltons and a pI of about 8.65. Clone DNA23334-1392 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

- 35 Analyzing the amino acid sequence of SEQ ID NO:14, the putative transmembrane domains are at about amino acids 16-36, 50-74, 147-168, 229-250, 271-293, 298-318 and 328-368 of SEQ ID NO:14. N-glycosylation sites are at about amino acids 128-131, 204-207, 218-221 and 274-377 of SEQ ID NO:14. The corresponding nucleotides can be routinely determined given the sequences provided herein.

PRO190 has sequence identity with at least the following Dayhoff sequences designated as: CEZK896_2, JC5023, GMS1_SCHPO and S44668.

EXAMPLE 8: Isolation of cDNA clones Encoding Human PRO341

A clone designated herein as DNA12920 was isolated as described in Example 2 above from a human placenta tissue library. The DNA12920 sequence is shown in Figure 13 (SEQ ID NO:21). The DNA12920 sequence was then compared to various EST databases including public EST databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify homologous EST sequences. The comparison was performed using the computer program BLAST or BLAST2 [Altschul et al., *Methods in Enzymology*, 266:460-480 (1996)]. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). This consensus sequence is herein designated DNA25314. Oligonucleotide primers based upon the DNA25314 sequence were then synthesized and employed to screen a human placenta cDNA library which resulted in the identification of the DNA26288-1239 clone shown in Figure 11. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 380-382, and a stop signal at nucleotide positions 1754-1756 (Figure 11, SEQ ID NO:19). The predicted polypeptide precursor is 458 amino acids long, has a calculated molecular weight of approximately 50,264 daltons and an estimated pI of approximately 8.17. Analysis of the full-length PRO341 sequence shown in Figure 12 (SEQ ID NO:20) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 17, transmembrane domains from about amino acid 171 to about amino acid 190, from about amino acid 220 to about amino acid 239, from about amino acid 259 to about amino acid 275, from about amino acid 286 to about amino acid 305, from about amino acid 316 to about amino acid 335, from about amino acid 353 to about amino acid 378 and from about amino acid 396 to about amino acid 417 and potential N-glycosylation sites from about amino acid 145 to about amino acid 147 and from about amino acid 155 to about amino acid 158. Clone DNA26288-1239 has been deposited with ATCC on April 21, 1998 and is assigned ATCC deposit no. 209792.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 12 (SEQ ID NO:20), evidenced homology between the PRO341 amino acid sequence and the following Dayhoff sequences: S75696, H69788, D69852, A69888, B64918, F64752, LPU89276_1, G64962, S52977 and S44253.

EXAMPLE 9: Isolation of cDNA clones Encoding Human PRO180

A clone designated herein as DNA12922 was isolated as described in Example 2 above from a human placenta tissue library. The DNA12922 sequence is shown in Figure 16 (SEQ ID NO:24). The DNA12922 sequence was then compared to various EST databases including public EST databases (e.g., GenBank), and a

proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify homologous EST sequences. The comparison was performed using the computer program BLAST or BLAST2 [Altschul et al., *Methods in Enzymology*, 266:460-480 (1996)]. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

- 5 An oligonucleotide probe was formed based upon the consensus sequence obtained above. This probe had the following sequence.

5'-ACCTGTTAGAAATGTGGTGGTTTCAGCAAGGCCTCAGTTT (SEQ ID NO:25).

- This probe was used to screen a human placenta library prepared as described in paragraph 1 of Example 2 above. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, 10 Holmes et al., *Science*, 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp. A clone designated herein as DNA26843-1389 was obtained.

- The entire nucleotide sequence of DNA26843-1389 is shown in Figure 14 (SEQ ID NO:22). Clone DNA26843-1389 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 121-123 and ending at the stop codon at nucleotide positions 919-921 (Figure 14). The predicted 15 polypeptide precursor is 266 amino acids long (Figure 15). The full-length PRO180 protein shown in Figure 15 has an estimated molecular weight of about 29,766 daltons and a pI of about 8.39. Clone DNA26843-1389 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

- Still analyzing the amino acid sequence of SEQ ID NO:23, the transmembrane domains are at about 20 amino acids 13-33 (type II), 54-73, 94-113, 160-180 and 122-141 of SEQ ID NO:23. N-myristoylation sites are at about amino acids 57-62, 95-100, 99-104, 124-129 and 183-188 of SEQ ID NO:23. The corresponding nucleotides can be routinely determined given the sequences provided herein.

- An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 15 (SEQ ID NO:23), evidenced some sequence 25 identity between the PRO180 amino acid sequence and the following Dayhoff sequences: CEC33A11_2, CEG11E6_5, CELW03A5_1 AND PEU83861_2 (NADH dehydrogenase subunit 4L, mitochondrion).

EXAMPLE 10: Isolation of cDNA clones Encoding Human PRO194

- A consensus DNA sequence was assembled relative to other EST sequences using phrap as described 30 in Example 1 above. This consensus sequence is herein DNA19464. Based on the DNA19464 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO194. PCR primers (forward and reverse) were synthesized based upon the DNA19464 sequence. Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA19464 sequence.

- 35 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO194 gene using the probe oligonucleotide and one of the PCR primers. RNA

for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB25).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO194 [herein designated as DNA26844-1394] (SEQ ID NO:27) and the derived protein sequence for PRO194.

The entire nucleotide sequence of DNA26844-1394 is shown in Figure 17 (SEQ ID NO:27). Clone
 5 DNA26844-1394 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 81-83 and ending at the stop codon at nucleotide positions 873-875 (Figure 17). The predicted polypeptide precursor is 264 amino acids long (Figure 18). The full-length PRO194 protein shown in Figure 18 has an estimated molecular weight of about 29,665 daltons and a pI of about 9.34. Analysis of the full-length PRO194 sequence shown in Figure 18 (SEQ ID NO:28) evidences the presence of various important
 10 polypeptides domains as shown in Figure 18. Clone DNA26844-1394 has been deposited with ATCC on June 2, 1998 and is assigned ATCC deposit no. 209926.

Analysis of the amino acid sequence of the full-length PRO194 polypeptide suggests that it does not exhibit significant sequence similarity to any known human protein. However, an analysis of the Dayhoff
 15 database (version 35.45 SwissProt 35) evidenced some homology between the PRO194 amino acid sequence and the following Dayhoff sequences, HUMORFT_1, CET07F10_5, ATFCA9_12, F64934, YDJX_ECOLI, ATAF00065719F29G20.19, H70002, S76980, H64934 and S76385.

EXAMPLE 11: Isolation of cDNA clones Encoding Human PRO203

A clone designated herein as DNA15618 was isolated as described in Example 2 above from a human
 20 fetal lung tissue library. The DNA15618 sequence is shown in Figure 21 (SEQ ID NO:31). Oligonucleotide probes were generated from the sequence of the DNA15618 molecule and were used to screen a human fetal lung library (LIB26) prepared as described in paragraph 1 of Example 2 above. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, **253**:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp.

A full length clone was identified that contained a single open reading frame with an apparent
 25 translational initiation site at nucleotide positions 159-161 and ending at the stop codon found at nucleotide positions 1200-1202 (Figure 19; SEQ ID NO:29). The predicted polypeptide precursor is 347 amino acids long, has a calculated molecular weight of approximately 39,870 daltons and an estimated pI of approximately 6.76. Analysis of the full-length PRO203 sequence shown in Figure 20 (SEQ ID NO:30) evidences the presence of
 30 the following: a type II transmembrane domain at about amino acid 64 to about amino acid 87; possible N-glycosylation sites at about amino acid 147 to about amino acid 150, about amino acid 155 to about amino acid 158, and about amino acid 237 to about amino acid 240; sequence identity with heavy-metal-associated domain proteins at about amino acid 23 to about amino acid 45, and sequence identity with D-isomer specific 2-hydroxyacid dehydrogenase at about amino acid 24 to about amino acid 34. Clone DNA30862-1396 was
 35 deposited with the ATCC on June 2, 1998, and is assigned ATCC deposit no. 209920.

Analysis of the amino acid sequence of the full-length PRO203 polypeptide suggests that it possesses sequence similarity to GST ATPase, thereby indicating that PRO203 may be a novel GST ATPase. More

specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO203 amino acid sequence and the following Dayhoff sequences, AF008124_1, CFRCD1GEN_1, and P_R82566.

EXAMPLE 12: Isolation of cDNA clones Encoding Human PRO290

5 An expressed sequence tag (EST) DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST was identified that had homology to beige and FAN. An oligonucleotide probe based upon the identified EST sequence was then synthesized and used to screen human fetal kidney cDNA libraries in an attempt to identify a full-length cDNA clone. The oligonucleotide probe had the following sequence:
5' TGACTGCACTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGCTG 3' (SEQ ID NO:34).

10 RNA for construction of cDNA libraries was isolated from human fetal kidney tissue. The cDNA libraries used to isolate the cDNA clones encoding human PRO290 were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as
15 pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science* 253:1278-1280 (1991)) in the unique XhoI and NotI.

A cDNA clone was identified and sequenced in entirety. The entire nucleotide sequence of DNA35680-1212 is shown in Figure 22 (SEQ ID NO:32). Clone DNA35680-1212 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 293-295, and a stop codon at nucleotide positions
20 3302-3304 (Figure 22; SEQ ID NO:32). The predicted polypeptide precursor is 1003 amino acids long.

It is currently believed that the PRO290 polypeptide is related to FAN and/or beige. Clone DNA35680-1212 has been deposited with ATCC and is assigned ATCC deposit no. 209790. It is understood that the deposited clone has the actual correct sequence rather than the representations provided herein. The full-length PRO290 protein shown in Figure 23 has an estimated molecular weight of about 112,013 daltons and a pI of
25 about 6.4.

EXAMPLE 13: Isolation of cDNA Clones Encoding Human PRO874

A consensus DNA sequence designated herein as DNA36459 was identified using phrap as described in Example 1 above. Based on the DNA36459 consensus sequence, oligonucleotides were synthesized: 1) to
30 identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the coding sequence for PRO874.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TCGTGCCCAGGGGCTGATGTGC-3' (SEQ ID NO:37); and

reverse PCR primer 5'-GTCTTTACCCAGCCCCGGGATGCG-3' (SEQ ID NO:38).

35 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA36459 sequence which had the following nucleotide sequence:

hybridization probe

5'-GGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGG-3' (SEQ ID NO:39).

In order to screen several libraries for a source of a clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO874 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB25).

DNA sequencing of the clones isolated as described above gave the DNA sequence for PRO874 [herein designated as DNA40621-1440] (SEQ ID NO:35) and the derived protein sequence for PRO874.

The entire nucleotide sequence of DNA40621-1440 is shown in Figure 24 (SEQ ID NO:35). Clone DNA40621-1440 contains a single open reading frame ending at the stop codon at nucleotide positions 964-966 (Figure 24). The predicted polypeptide encoded by DNA40621-1440 is 321 amino acids long (Figure 25). The PRO874 protein shown in Figure 25 has an estimated molecular weight of about 36,194 daltons and a pI of about 9.85. Analysis of the PRO874 sequence shown in Figure 25 (SEQ ID NO:36) evidenced the presence of the following: a type II transmembrane domain at about amino acids 57-80; additional transmembrane domains at about amino acids 110-126, 215-231, and 254-274; potential N-glycosylation sites at about amino acids 16-19, 27-30, and 289-292; sequence identity with hypothetical YBR002c family proteins at about amino acids 276-287; and sequence identity with ammonium transporter proteins at about amino acids 204-230. Clone DNA40621-1440 was deposited with the ATCC on June 2, 1998, and is assigned ATCC deposit no. 209922.

Analysis of the amino acid sequence of the PRO874 polypeptide suggests that it is a novel multi-span transmembrane protein. However, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced sequence identity between the PRO874 amino acid sequence and the following Dayhoff sequences: S67049, AF054839_1, S73437, S52460, and HIVU80570_1.

EXAMPLE 14: Isolation of cDNA Clones Encoding Human PRO710

A yeast screening assay was employed to identify cDNA clones that encoded potential secreted proteins. Use of this yeast screening assay allowed identification of a single cDNA clone whose sequence (herein designated as DNA38190) is shown in Figure 28 (SEQ ID NO:42). Based on the DNA38190 sequence shown in Figure 28, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO710. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TTCCGCAAAGAGTTCTACGAGGTGG-3' (SEQ ID NO:43)

reverse PCR primer 5'-ATTGACAACATTGACTGGCCTATGGG-3' (SEQ ID NO:44)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA38190 sequence which had the following nucleotide sequence

hybridization probe

5'-GTGGATGCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTCCAGTGTGA-3' (SEQ ID NO:45)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO710 gene using the probe oligonucleotide and one of the PCR primers.

5 RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 10 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 67-69 and ending at the stop codon found at nucleotide positions 1765-1767 (Figure 26, SEQ ID NO:40). The predicted polypeptide precursor is 566 amino acids long, has a 15 calculated molecular weight of approximately 65,555 daltons and an estimated pI of approximately 5.44. Analysis of the full-length PRO710 sequence shown in Figure 27 (SEQ ID NO:41) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 32, a transmembrane domain from about amino acid 454 to about amino acid 476, an aminoacyl-transfer RNA synthetase class-II signature sequence from about amino acid 6 to about amino acid 26 and potential N-glycosylation sites from about amino acid 111 20 to about amino acid 114, from about amino acid 146 to about amino acid 149 and from about amino acid 292 to about amino acid 295. Clone DNA44161-1434 has been deposited with ATCC on May 27, 1998 and is assigned ATCC deposit no. 209907.

Analysis of the amino acid sequence of the full-length PRO710 polypeptide suggests that it possesses significant sequence similarity to the CDC45 protein, thereby indicating that PRO710 may be a novel CDC45 25 homolog. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO710 amino acid sequence and the following Dayhoff sequences, HSAJ3728_1, CEF34D10_1, S64939, UMUS0276_1, TRHY_SHEEP, CELT14E8_1, RNA1_YEAST, LVU89340_1, HSU80736_1 and CEZK337_2.

30 EXAMPLE 15: Isolation of cDNA clones Encoding Human PRO1151

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA40665. Based on the DNA40665 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for 35 PRO1151.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCAGACGCTGCTCTTCGAAAGGGTC-3' (SEQ ID NO:48)

reverse PCR primer 5'-GGTCCCCGTAGGCCAGGTCCAGC-3' (SEQ ID NO:49)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA40665 sequence which had the following nucleotide sequence

5 hybridization probe

5'-CTACTTCTTCAGCCTCAATGTGCACAGCTGGAATTACAAGGAGACGTACG-3' (SEQ ID NO:50)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1151 gene using the probe oligonucleotide and one of the PCR primers. RNA
10 for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1151 (designated herein as DNA44694-1500 [Figure 29, SEQ ID NO:46]; and the derived protein sequence for PRO1151.

The entire nucleotide sequence of DNA44694-1500 is shown in Figure 29 (SEQ ID NO:46). Clone
15 DNA44694-1500 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 272-274 and ending at the stop codon at nucleotide positions 1049-1051 (Figure 29). The predicted polypeptide precursor is 259 amino acids long (Figure 30). The full-length PRO1151 protein shown in Figure 30 has an estimated molecular weight of about 28,770 daltons and a pI of about 6.12. Analysis of the full-length PRO1151 sequence shown in Figure 30 (SEQ ID NO:47) evidences the presence of the following: a signal
20 peptide from about amino acid 1 to about amino acid 20, a potential N-glycosylation site from about amino acid 72 to about amino acid 75 and amino acid sequence blocks having homology to C1q domain-containing proteins from about amino acid 144 to about amino acid 178, from about amino acid 78 to about amino acid 111 and from about amino acid 84 to about amino acid 117. Clone UNQ581 (DNA44694-1500) has been deposited with ATCC on August 11, 1998 and is assigned ATCC deposit no. 203114.

25 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 30 (SEQ ID NO:47), evidenced significant homology between the PRO1151 amino acid sequence and the following Dayhoff sequences: ACR3_HUMAN, HP25_TAMAS, HUMC1QB2_1, P_R99306, CA1F_HUMAN, JX0369, CA24_HUMAN, S32436, P_R28916 and CA54_HUMAN.

30

EXAMPLE 16: Isolation of cDNA clones Encoding Human PRO1282

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein as DNA33778. Based on the DNA33778 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained
35 the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1282.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'TCTTCAGCCGCTTGCGCAACCTC3' (SEQ ID NO:53); and

reverse PCR primer 5'TTGCTCACATCCAGCTCCTGCAGG3' (SEQ ID NO:54).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA33778 sequence which had the following nucleotide sequence:

5 hybridization probe

5'TGGATGTTGTCCAGACAACCAGCTGGAGCTGTATCCGAGGC3' (SEQ ID NO:55).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1282 gene using the probe oligonucleotide and one of the PCR primers. RNA
10 for construction of the cDNA libraries was isolated from human fetal liver.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1282 (designated herein as DNA45495-1550 [Figure 31, SEQ ID NO:51]; and the derived protein sequence for PRO1282.

The entire coding sequence of PRO1282 is shown in Figure 31 (SEQ ID NO:51). Clone DNA45495-
15 1550 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 120-122, and an apparent stop codon at nucleotide positions 2139-2141 (SEQ ID NO:51). The predicted polypeptide precursor is 673 amino acids long. The signal peptide is at about amino acids 1-23; the transmembrane domain is at about amino acids 579-599; an EGF-like domain cysteine pattern signature starts at about amino acid 430; and leucine zipper patterns start at about amino acids 197 and 269 of SEQ ID NO:52,
20 see Figure 32. Clone DNA45495-1550 has been deposited with the ATCC and is assigned ATCC deposit no. 203156. The full-length PRO1282 protein shown in Figure 32 has an estimated molecular weight of about 71,655 daltons and a pI of about 7.8.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 32 (SEQ ID NO:52), revealed sequence identity
25 between the PRO1282 amino acid sequence and the following Dayhoff sequences (data from database incorporated by reference): AB007876_1, RNPLGPV_1, MUSLRRP_1, ALS_PAPPA, AC004142_1, ALS_HUMAN, AB014462_1, DMTARTAN_1, HSCHON03_1 and S46224.

EXAMPLE 17: Isolation of cDNA clones Encoding Human PRO358

30 Using the method described in Example 1 above, a single EST sequence was identified in the Incyte database, designated herein as INC3115949. Based on the INC3115949 EST sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO358.

A pair of PCR primers (forward and reverse) were synthesized:

35 forward PCR primer 5'-TCCCACCAGGTATCATAAACTGAA-3' (SEQ ID NO:58)

reverse PCR primer 5'-TTATAGACAATCTGTTCTCATCAGAGA-3' (SEQ ID NO:59)

A probe was also synthesized:

5'-AAAAAGCATACTTGAATGGCCCAAGGATAGGTGTAAATG-3' (SEQ ID NO:60)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO358 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human bone marrow (LIB256). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO358 (Figure 33, SEQ ID NO:56) and the derived protein sequence for PRO358 (Figures 34, SEQ ID NO:57).

The entire nucleotide sequence of the clone identified (DNA47361-1154) is shown in Figure 33 (SEQ ID NO:56). Clone DNA47361-1154 contains a single open reading frame with an apparent translational initiation site (ATG start signal) at nucleotide positions underlined in Figure 33. The predicted polypeptide precursor is 811 amino acids long, including a putative signal sequence (amino acids 1 to 19), an extracellular domain (amino acids 20 to 575, including leucine rich repeats in the region from position 55 to position 575), a putative transmembrane domain (amino acids 576 to 595). Clone DNA47361-1249 has been deposited with ATCC and is assigned ATCC deposit no. 209431.

EXAMPLE 18: Isolation of cDNA clones Encoding Human PRO1310

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein as DNA37164. Based on the DNA37164 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1310.

PCR primers (forward and reverse) were synthesized:

forward PCR primer: 5'GTTCTCAATGAGCTACCCGTCCCC3' (SEQ ID NO:63) and
reverse PCR primer: 5'CGCGATGTAGTGGAAGCTCGGGCTC3' (SEQ ID NO:64).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA47394 sequence which had the following nucleotide sequence:

hybridization probe:

5'ATCCGCATAAACCCCTCAGTCCTGGTTTGATAATGGGAGCATCTGCATGAG3' (SEQ ID NO:65).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to

isolate clones encoding the PRO1310 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1310 and the derived protein sequence for PRO1310.

The entire coding sequence of PRO1310 is shown in Figures 35A-B (SEQ ID NO:61). Clone DNA47394-1572 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 326-328, and an apparent stop codon at nucleotide positions 2594-2596 (SEQ ID NO:61). The predicted polypeptide precursor is 765 amino acids long. The signal peptide is at about amino acids 1-25 of SEQ ID NO:62. Clone DNA47394-1572 has been deposited with ATCC and is assigned ATCC deposit no. 203109. The full-length PRO1310 protein shown in Figure 36 has an estimated molecular weight of about 85,898 daltons and a pI of about 6.87.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 36 (SEQ ID NO:62), revealed sequence identity between the PRO1310 amino acid sequence and the following Dayhoff sequences: AF017639_1, P_W36817, JC5256, CBPH_HUMAN, MMU23184_1, CBPN_HUMAN, HSU83411_1, CEF01D4_7, RNU62897_1 and P_W11851.

EXAMPLE 19: Isolation of cDNA Clones Encoding Human PRO698

A yeast screening assay was employed to identify cDNA clones that encoded potential secreted proteins. Use of this yeast screening assay allowed identification of a single cDNA clone whose sequence (herein designated as DNA39906) is shown in Figure 39 (SEQ ID NO:68). Based on the DNA39906 sequence shown in Figure 39, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO698. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AGCTGTGGTCATGGTGGTGTGGTG-3' (SEQ ID NO:69)

reverse PCR primer 5'-CTACCTTGCCATAGGTGATCCGC-3' (SEQ ID NO:70)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA39906 sequence which had the following nucleotide sequence

hybridization probe

5'-CATCAGCAAACCGTCTGTGGTTCAGCTCAACTGGAGAGGGTT-3' (SEQ ID NO:71)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO698 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human bone marrow tissue (LIB255). The cDNA

libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to Sall hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 14-16 and ending at the stop codon found at nucleotide positions 1544-1546 (Figure 37, SEQ ID NO:66). The predicted polypeptide precursor is 510 amino acids long, has a calculated molecular weight of approximately 57,280 daltons and an estimated pI of approximately 5.61.

Analysis of the full-length PRO698 sequence shown in Figure 38 (SEQ ID NO:67) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 20, potential N-glycosylation sites from about amino acid 72 to about amino acid 75, from about amino acid 136 to about amino acid 139, from about amino acid 193 to about amino acid 196, from about amino acid 253 to about amino acid 256, from about amino acid 352 to about amino acid 355 and from about amino acid 411 to about amino acid 414 an amino acid block having homology to legume lectin beta-chain proteins from about amino acid 20 to about amino acid 39 and an amino acid block having homology to the HBGF/FGF family of proteins from about amino acid 338 to about amino acid 365. Clone DNA48320-1433 has been deposited with ATCC on May 27, 1998 and is assigned ATCC deposit no. 209904.

Analysis of the amino acid sequence of the full-length PRO698 polypeptide suggests that it possesses significant sequence similarity to the olfactomedin protein, thereby indicating that PRO698 may be a novel olfactomedin homolog. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO698 amino acid sequence and the following Dayhoff sequences, OLFM_RANCA, I73637, AB006686S3_1, RNU78105_1, RNU72487_1, P_R98225, CELC48E7_4, CEF11C3_3, XLU85970_1 and S42257.

EXAMPLE 20: Isolation of cDNA Clones Encoding Human PRO732

A yeast screening assay was employed to identify cDNA clones that encoded potential secreted proteins. Use of this yeast screening assay allowed identification of a single cDNA clone whose sequence (herein designated as DNA42580) is shown in Figure 45 (SEQ ID NO:77). The DNA42580 sequence was then compared to a variety of known EST sequences to identify homologies. The EST databases employed included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

Using the above analysis, a consensus DNA sequence was assembled relative to other EST sequences using phrap. This consensus sequence is herein designated consen01. Proprietary Genentech EST sequences were employed in the consensus assembly and they are herein designated DNA20239 (Figure 42; SEQ ID NO:74), DNA38050 (Figure 43; SEQ ID NO:75) and DNA40683 (Figure 44; SEQ ID NO:76).

Based on the consen01 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO732. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-ATGTTTGTGTGGAAGTGCCCCG-3' (SEQ ID NO:78)

forward PCR primer 5'-GTCAACATGCTCCTCTGC-3' (SEQ ID NO:79)

reverse PCR primer 5'-AATCCATTGTGCACTGCAGCTCTAGG-3' (SEQ ID NO:80)

reverse PCR primer 5'-GAGCATGCCACCACTGGACTGAC-3' (SEQ ID NO:81)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA44143 sequence which had the following nucleotide sequence

hybridization probe

5'-GCCGATGCTGTCCTAGTGGAACAACCTCCACTGTAAGTAGATTGATCTATGCAC-3' (SEQ ID NO:82)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO732 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB26). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 88-90 and ending at the stop codon found at nucleotide positions 1447-1449 (Figure 40, SEQ ID NO:72). The predicted polypeptide precursor is 453 amino acids long, has a calculated molecular weight of approximately 50,419 daltons and an estimated pI of approximately 5.78. Analysis of the full-length PRO732 sequence shown in Figure 41 (SEQ ID NO:73) evidences the presence of

the following: a signal peptide from about amino acid 1 to about amino acid 28, transmembrane domains from about amino acid 37 to about amino acid 57, from about amino acid 93 to about amino acid 109, from about amino acid 126 to about amino acid 148, from about amino acid 151 to about amino acid 172, from about amino acid 197 to about amino acid 215, from about amino acid 231 to about amino acid 245, from about amino acid 260 to about amino acid 279, from about amino acid 315 to about amino acid 333, from about amino acid 384 to about amino acid 403 and from about amino acid 422 to about amino acid 447, potential N-glycosylation sites from about amino acid 33 to about amino acid 36, from about amino acid 34 to about amino acid 37, from about amino acid 179 to about amino acid 183, from about amino acid 298 to about amino acid 301, from about amino acid 337 to about amino acid 340 and from about amino acid 406 to about amino acid 409, an amino acid block having homology to the MIP family of proteins from about amino acid 119 to about amino acid 149 and an amino acid block having homology to DNA/RNA non-specific endonuclease proteins from about amino acid 279 to about amino acid 286. Clone DNA48334-1435 has been deposited with ATCC on June 2, 1998 and is assigned ATCC deposit no. 209924.

Analysis of the amino acid sequence of the full-length PRO732 polypeptide suggests that it possesses significant sequence similarity to the Diff33 protein, thereby indicating that PRO732 may be a novel Diff33 homolog. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO732 amino acid sequence and the following Dayhoff sequences, HS179M20_2, MUSTETU_1, CER11H6_2, RATDRP_1, S51256, E69226, AE000869_1, JC4120, CYB_PARTE and P_R50619.

20 EXAMPLE 21: Isolation of cDNA clones Encoding Human PRO1120

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein consen0352. The consen0352 sequence was then extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above. The extended consensus sequence is designated herein as DNA34365. Based on the DNA34365 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1120.

PCR primers (forward and reverse) were synthesized:

forward PCR primers: 5'-GAAGCCGGCTGTCTGAATC-3' (SEQ ID NO:85),

5'-GGCCAGCTATCTCCGCAG-3' (SEQ ID NO:86), 5'-AAGGGCCTGCAAGAGAAG-3' (SEQ ID NO:87),

5'-CACTGGGACAACTGTGGG-3' (SEQ ID NO:88),

5'-CAGAGGCAACGTGGAGAG-3' (SEQ ID NO:89), and

5'-AAGTATTGTCATACAGTGTTTC-3' (SEQ ID NO:90);

reverse PCR primers: 5'-TAGTACTTGGGCACGAGGTTGGAG-3' (SEQ ID NO:91), and 5'-

TCATACCAACTGCTGGTCATTGGC-3' (SEQ ID NO:92).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA34365 consensus sequence which had the following nucleotide sequence:

hybridization probe:

5'-CTCAAGCTGCTGGACACGGAGCGGCCGGTGAATCGGTTTCACTTG-3' (SEQ ID NO:93).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO1120 gene using the probe oligonucleotide and one of the PCR primers. RNA
5 for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1120 (designated herein as DNA48606-1479 [Figures 46A-B, SEQ ID NO:83]; and the derived protein sequence for PRO1120.

The entire coding sequence of PRO1120 is shown in Figures 46A-B (SEQ ID NO:83). Clone
10 DNA48606-1479 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 608-610 and an apparent stop codon at nucleotide positions 3209-3211. The predicted polypeptide precursor is 867 amino acids long. The full-length PRO1120 protein shown in Figure 47 has an estimated molecular weight of about 100,156 Daltons and a pI of about 9.44. Additional features of the PRO1120 polypeptide include a signal peptide at about amino acids 1-17; a sulfatase signature at about amino acids 86-98;
15 regions of homology to sulfatases at about amino acids 87-106, 133-146, 216-229, 291-320, and 365-375; and potential N-glycosylation sites at about amino acids 65-68, 112-115, 132-135, 149-152, 171-174, 198-201, 241-245, 561-564, 608-611, 717-720, 754-757, and 764-767.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 47 (SEQ ID NO:84), revealed significant
20 homology between the PRO1120 amino acid sequence and the following Dayhoff sequences: CELK09C4_1, GL6S_HUMAN, G65169, NCU89492_1, BCU44852_1, E64903, P_R51355, STS_HUMAN, GA6S_HUMAN, and IDS_MOUSE. Clone DNA48606-1479 was deposited with the ATCC on July 1, 1998, and is assigned ATCC deposit no. 203040.

25 EXAMPLE 22: Isolation of cDNA clones Encoding Human PRO537

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated as Incyte EST cluster no. 29605. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto,
30 CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA48350.

35 In light of an observed sequence homology between the DNA48350 consensus sequence and an EST sequence encompassed within the Merck EST clone no. R63443, the Merck EST clone R63443 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.

The sequence of this cDNA insert is shown in Figure 48 and is herein designated as DNA49141-1431.

Clone DNA49141-1431 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 97-99 and ending at the stop codon at nucleotide positions 442-444 (Figure 48). The predicted polypeptide precursor is 115 amino acids long (Figure 49). The full-length PRO537 protein shown in Figure 49 has an estimated molecular weight of about 13,183 daltons and a pI of about 12.13. Analysis of the full-length PRO537 sequence shown in Figure 49 (SEQ ID NO:95) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 31, a potential N-glycosylation site from about amino acid 44 to about amino acid 47, potential N-myristoylation sites from about amino acid 3 to about amino acid 8 and from about amino acid 16 to about amino acid 21 and an amino acid block having homology to multicopper oxidase proteins from about amino acid 97 to about amino acid 105. Clone DNA49141-1431 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203003.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 49 (SEQ ID NO:95), evidenced homology between the PRO537 amino acid sequence and the following Dayhoff sequences: A54523, CELF22H10_2, FK44_MOUSE, OTX1_HUMAN, URB1_USTMA, KNOB_PLAFN, A32895_1, AF036332_1, HRG_HUMAN and HRP3_PLAFS.

EXAMPLE 23: Isolation of cDNA clones Encoding Human PRO536

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated herein as ss.clu2437.init. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA48351.

In light of an observed sequence homology between the DNA48351 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H11129, the Merck EST clone H11129 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 50 and is herein designated as DNA49142-1430.

Clone DNA49142-1430 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 48-50 and ending at the stop codon at nucleotide positions 987-989 (Figure 50). The predicted polypeptide precursor is 313 amino acids long (Figure 51). The full-length PRO536 protein shown in Figure 51 has an estimated molecular weight of about 34,189 daltons and a pI of about 4.8. Analysis of the full-length PRO536 sequence shown in Figure 51 (SEQ ID NO:97) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 25, a potential N-glycosylation site from about amino acid 45 to about amino acid 48 and an amino acid sequence block having homology to sulfatase proteins from

about amino acid 16 to about amino acid 26. Clone DNA49142-1430 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203002.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 51 (SEQ ID NO:97), evidenced homology between the PRO536 amino acid sequence and the following Dayhoff sequences: APU46857_1, PK2_DICDI, H64743, F5114_18, CEAM_ECOLI, GEN14267, H64965, TCU39815_1, PSBJ_ODOSI and P_R06980.

EXAMPLE 24: Isolation of cDNA clones Encoding Human PRO535

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated herein as ss.clu12694.init. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA48352. Two proprietary Genentech EST sequences were employed in the assembly are herein shown in Figures 54 and 55.

In light of an observed sequence homology between the DNA48352 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H86994, the Merck EST clone H86994 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 52 and is herein designated as DNA49143-1429.

Clone DNA49143-1429 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 78-80 and ending at the stop codon at nucleotide positions 681-683 (Figure 52). The predicted polypeptide precursor is 201 amino acids long (Figure 53). The full-length PRO535 protein shown in Figure 53 has an estimated molecular weight of about 22,180 daltons and a pI of about 9.68. Analysis of the full-length PRO535 sequence shown in Figure 53 (SEQ ID NO:99) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 25, a transmembrane domain from about amino acid 155 to about amino acid 174, a potential N-glycosylation site from about amino acid 196 to about amino acid 199 and FKBP-type peptidyl-prolyl cis-trans isomer signature sequences from about amino acid 62 to about amino acid 77, from about amino acid 87 to about amino acid 123 and from about amino acid 128 to about amino acid 141. Clone DNA49143-1429 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203013.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST- sequence alignment analysis of the full-length sequence shown in Figure 53 (SEQ ID NO:99), evidenced homology between the PRO535 amino acid sequence and the following Dayhoff sequences: S71237, P_R93551, P_R28980, S71238, FKBP2_HUMAN, CELC05C8_1, S55383, S72485, CELC50F2_6 and S75144.

EXAMPLE 25: Isolation of cDNA clones Encoding Human PRO718

A cDNA sequence isolated in the amylase screen described in Example 2 (human fetal lung library) above is herein designated DNA43512 (see Figure 62; SEQ ID NO:108). The DNA43512 sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA45625. Proprietary Genentech EST sequences were employed in the assembly and are herein shown in Figures 58-61.

Based on the DNA45625 sequence, oligonucleotide probes were generated and used to screen a human fetal lung library (LIB25) prepared as described in paragraph 1 of Example 2 above. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GGGTGGATGGTACTGCTGCATCC-3' (SEQ ID NO:109)

reverse PCR primer 5'-TGTTGTGCTGTGGGAAATCAGATGTG-3' (SEQ ID NO:110)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA45625 sequence which had the following nucleotide sequence:

hybridization probe
5'-GTGCTCTGGAGGCTGTGGCCGTTTGTCTTCTGGGCTAAAATCGGG-3' (SEQ ID NO:111)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO718 gene using the probe oligonucleotide and one of the PCR primers.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 36-38 and ending at the stop codon found at nucleotide positions 607-609 (Figure 56; SEQ ID NO:102). The predicted polypeptide precursor is 157 amino acids long, has a calculated molecular weight of approximately 17,400 daltons and an estimated pI of approximately 5.78. Analysis of the full-length PRO718 sequence shown in Figure 57 (SEQ ID NO:103) evidences the presence of the following: a type II transmembrane domain from about amino acid 21 to about amino acid 40, and other transmembrane domains at about amino acid 58 to about amino acid 78, about amino acid 95 to about amino acid 114, and about amino acid 127 to about amino acid 147; a cell attachment sequence from about amino acid 79 to about amino acid 81; and a potential N-glycosylation site from about amino acid 53 to about amino acid 56. Clone DNA49647-1398 has been deposited with ATCC on June 2, 1998 and is assigned ATCC deposit no. 209919.

Analysis of the amino acid sequence of the full-length PRO718 polypeptide suggests that it possesses no significant sequence similarity to any known protein. However, an analysis of the Dayhoff database (version

35.45 SwissProt 35) evidenced some degree of homology between the PRO718 amino acid sequence and the following Dayhoff sequences: AF045606_1, AF039906_1, SPBC8D2_2, S63441, F64728, COX1_TRYBB, F64375, E64173, RPYGJT_3, MTCY261_23.

EXAMPLE 26: Isolation of cDNA clones Encoding Human PRO872

5 Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST sequence designated herein as clu120709.init. The clu120709.init sequence was then compared a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or
10 in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA48254.

In light of an observed sequence homology between the DNA48254 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3438068, the Incyte EST clone 3438068 was purchased
15 and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 63 and is the full-length DNA sequence for PRO872. Clone DNA49819-1439 was deposited with the ATCC on June 2, 1998, and is assigned ATCC deposit no. 209931.

The entire nucleotide sequence of DNA49819-1439 is shown in Figure 63 (SEQ ID NO:112). Clone
20 DNA49819-1439 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 14-16 and ending at the stop codon at nucleotide positions 1844-1846 (Figure 63). The predicted polypeptide precursor is 610 amino acids long (Figure 64). The full-length PRO872 protein shown in Figure 64 has an estimated molecular weight of about 66,820 daltons and a pI of about 8.65. Analysis of the full-length PRO872 sequence shown in Figure 64 (SEQ ID NO:113) evidences the presence of the following features: a
25 signal peptide at amino acid 1 to about 18, putative transmembrane domains at about amino acids 70-87, 200-222 and 568-588; sequence identity with bacterial-type phytoene dehydrogenase protein at about amino acids 71-105; sequence identity with a regulator of chromosome condensation (RCC1) signature 2 at about amino acids 201-211; leucine zipper patterns at about amino acids 214-235, 221-242, 228-249 and 364-385; a potential N-glycosylation site at about amino acids 271-274; and a glycosaminoglycan attachment site at about amino acids
30 75-78. Analysis of the amino acid sequence of the full-length PRO872 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO872 amino acid sequence and the following Dayhoff sequences: PRCRTI_1, S75951, S74689, CELF37C4_3, CRTI_RHOCA, S76617, YNI2_METTL, MTV014_14, AOFB_HUMAN, and MMU70429_1.

35 EXAMPLE 27: Isolation of cDNA clones Encoding Human PRO1063

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence designated herein as ss.clu119743.init. The Incyte EST cluster sequence

ss.clu119743.init sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA48288.

In light of an observed sequence homology between the DNA48288 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2783726, the Incyte EST clone 2783726 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 65 and is herein designated DNA49820-1427.

The full length clone shown in Figure 65 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 90-92 and ending at the stop codon found at nucleotide positions 993-995 (Figure 65; SEQ ID NO:114). The predicted polypeptide precursor is 301 amino acids long, has a calculated molecular weight of approximately 33,530 daltons and an estimated pI of approximately 4.80. Analysis of the full-length PRO1063 sequence shown in Figure 66 (SEQ ID NO:115) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 21, potential N-glycosylation sites from about amino acid 195 to about amino acid 198, from about amino acid 217 to about amino acid 220 and from about amino acid 272 to about amino acid 275, a glycosaminoglycan attachment site from about amino acid 267 to about amino acid 270, a microbodies C-terminal targeting signal site from about amino acid 299 to about amino acid 301, a type II fibronectin collagen-binding domain homology sequence from about amino acid 127 to about amino acid 168 and a fructose-bisphosphate aldolase class II protein homology sequence from about amino acid 101 to about amino acid 118. Clone DNA49820-1427 has been deposited with the ATCC on June 2, 1998 and is assigned ATCC deposit no. 209932.

Analysis of the amino acid sequence of the full-length PRO1063 polypeptide suggests that it possesses sequence similarity to the human type IV collagenase protein. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some degree of homology between the PRO1063 amino acid sequence and the following Dayhoff sequences, S68303, CFU68533_1, P_P91139, RNU65656_1, PA2R_RABIT, MMU56734_1, FINC_XENLA, A48925, P_R92778 and FA12_HUMAN.

EXAMPLE 28: Isolation of cDNA clones Encoding Human PRO619

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated herein as 88434. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score

of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

In light of an observed sequence homology between the consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1656694, the Incyte EST clone 1656694 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 67 and is herein designated as DNA49821-1562.

The full length clone shown in Figure 67 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 81-83 and ending at the stop codon found at nucleotide positions 450-452 (Figure 67; SEQ ID NO:116). The predicted polypeptide precursor (Figure 68, SEQ ID NO:117) is 123 amino acids long including a predicted signal peptide at about amino acids 1-20. PRO619 has a calculated molecular weight of approximately 13,710 daltons and an estimated pI of approximately 5.19. Clone DNA49821-1562 was deposited with the ATCC on June 16, 1998 and is assigned ATCC deposit no. 209981.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 68 (SEQ ID NO:117), revealed significant homology between the PRO619 amino acid sequence and the following Dayhoff sequences: S35302, D87009_1, HSU93494_1, HUMIGLAM5_1, D86999_2, HUMIGLYM1_1, HUMIGLYMKE_1, A29491_1, A29498_1, and VPR2_MOUSE.

EXAMPLE 29: Isolation of cDNA clones Encoding Human PRO943

A consensus DNA sequence encoding PRO943 was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence was then extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above. The extended consensus sequence is herein designated DNA36360. Based on the DNA36360 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO943.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CGAGATGACGCCGAGCCCCC-3' (SEQ ID NO:120)

reverse PCR primer 5'-CGGTTTCGACACGCGGCAGGTG-3' (SEQ ID NO:121)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA36360 sequence which had the following nucleotide sequence

hybridization probe

5'-TGCTGCTCCTGCTGCCGCCGCTGCTGCTGGGGGCCTTCCCGCCGG-3' (SEQ ID NO:122)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO943 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO943 (designated herein as DNA52192-1369 [Figure 69, SEQ ID NO:118]) and the derived protein sequence for PRO943.

The entire nucleotide sequence of DNA52192-1369 is shown in Figure 69 (SEQ ID NO:118). Clone DNA52192-1369 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 150-152 and ending at the stop codon at nucleotide positions 1662-1664 (Figure 69). The predicted polypeptide precursor is 504 amino acids long (Figure 70). The full-length PRO943 protein shown in Figure 70 has an estimated molecular weight of about 54,537 daltons and a pI of about 10.04. Analysis of the full-length PRO943 sequence shown in Figure 70 (SEQ ID NO:119) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 17, a transmembrane domain from about amino acid 376 to about amino acid 396, tyrosine kinase phosphorylation sites from about amino acid 212 to about amino acid 219 and from about amino acid 329 to about amino acid 336, potential N-glycosylation sites from about amino acid 111 to about amino acid 114, from about amino acid 231 to about amino acid 234, from about amino acid 255 to about amino acid 258 and from about amino acid 293 to about amino acid 296 and an immunoglobulin and MHC protein sequence homology block from about amino acid 219 to about amino acid 236. Clone DNA52192-1369 has been deposited with ATCC on July 1, 1998 and is assigned ATCC deposit no. 203042.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 70 (SEQ ID NO:119), evidenced significant homology between the PRO943 amino acid sequence and the following Dayhoff sequences: B49151, A39752, FGR1_XENLA, S38579, RATHBFGFRB_1, TVHU2F, FGR2_MOUSE, CEK3_CHICK, P_R21080 and A27171_1.

EXAMPLE 30: Isolation of cDNA clones Encoding Human PRO1188

A consensus DNA sequence was assembled relative to other EST sequences using the program "phrap" as described in Example 1 above. This consensus sequence is designated herein as DNA45679. Based on the DNA45679 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1188.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CTGGTGCCTCAACAGGGAGCAG-3' (SEQ ID NO:125)
reverse PCR primer 5'-CCATTGTGCAGGTCAGGTCACAG-3' (SEQ ID NO:126)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA45679 sequence which had the following nucleotide sequence:

hybridization probe

5'-CTGGAGCAAGTGCTCAGCTGCCTGTGGTCAGACTGGGGTC-3' (SEQ ID NO:127)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1188 gene using the probe oligonucleotide and one of the PCR primers. RNA

for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1188 (designated herein as DNA52598-1518 [Figure 71, SEQ ID NO:123]); and the derived protein sequence for PRO1188.

The entire coding sequence of PRO1188 is shown in Figure 71 (SEQ ID NO:123). Clone DNA52598-1518 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 136-138 and an apparent stop codon at nucleotide positions 3688-3690. The predicted polypeptide precursor is 1184 amino acids long. The full-length PRO1188 protein shown in Figure 72 has an estimated molecular weight of about 132,582 Daltons and a pI of about 8.80. Additional features include: a signal peptide at about amino acids 1-31; an ATP/GTP binding site motif A (P-loop) at about amino acids 266-273; an aldehyde dehydrogenases cysteine active site at about amino acids 188-199; growth factor and cytokines receptors family signature 2 at about amino acids 153-159; and potential N-glycosylation sites at about amino acids 129-132, 132-135, 346-349, 420-423, 550-553, 631-634, 1000-1003, and 1056-1059.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 72 (SEQ ID NO:124), revealed significant homology between the PRO1188 amino acid sequence and the following Dayhoff sequences: SSU83114_1, S56015, CET21B6_4, CELT19D2_1, and TSP1_MOUSE.

Clone DNA52598-1518 has been deposited with ATCC and is assigned ATCC deposit no 203107.

EXAMPLE 31: Isolation of cDNA clones Encoding Human PRO1133

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This sequence was extended using repeated cycles of phrap. The extended consensus sequence is designated herein DNA38102. Based on the DNA38102 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1133.

PCR primers (two forward and one reverse) were synthesized:
forward PCR primer 1 5'-TCGATTATGGACGAACATGGCAGC-3' (SEQ ID NO:130);
forward PCR primer 2 5'-TTCTGAGATCCCTCATCCTC-3' (SEQ ID NO:131); and
reverse primer 5'-AGGTCAGGGACAGCAAGTTTGGG-3' (SEQ ID NO:132).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA38102 sequence which had the following nucleotide sequence:

hybridization probe
 5'TTTGCTGGACCTCGGCTACGGAATTGGCTTCCCTCTACGGACAGCTGGAT3' (SEQ ID NO:133).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with a PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1133 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1133 and the derived protein sequence for PRO1133.

The entire coding sequence of PRO1133 is shown in Figure 73 (SEQ ID NO:128). Clone DNA53913-1490 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 266-268 and an apparent stop codon at nucleotide positions 1580-1582 of SEQ ID NO:128. The predicted polypeptide precursor is 438 amino acids long. The signal peptide is at amino acids 1-18 of SEQ ID NO:129. EGF-like domain cysteine pattern signatures start at 315 and 385 of SEQ ID NO:129 as shown in Figure 74. Clone DNA53913-1490 has been deposited with ATCC and is assigned ATCC deposit no. 203162. The full-length PRO1133 protein shown in Figure 74 has an estimated molecular weight of about 49,260 daltons and a pI of about 6.15.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 74 (SEQ ID NO:129), revealed some sequence identity between the PRO1133 amino acid sequence and the following Dayhoff sequences (data from the database incorporated herein): AF002717_1, LMG1_HUMAN, B54665, UNC6_CAEEL, LML1_CAEEL, M15_MOUSE, MMU88353_1, LMA1_HUMAN, HSLN2C64_1 and AF005258_1.

EXAMPLE 32: Isolation of cDNA clones Encoding Human PRO784

An initial DNA sequence (SEQ ID NO:136), referred to herein as DNA44661 and shown in Figure 77, was identified using a yeast screen, in a human fetal lung cDNA library that preferentially represents the 5' ends of the primary cDNA clones. DNA44661 was then compared to ESTs from public databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA), using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)]. The ESTs were then clustered and assembled into a consensus DNA sequence using the computer program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained is designated herein as "DNA45463". Based on the DNA45463 consensus sequence, oligonucleotides were synthesized for use as probes to isolate a clone of the full-length coding sequence for PRO784 from a human fetal lung cDNA library.

The full length DNA53978-1443 clone shown in Figure 75 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 37-39 and ending at the stop codon found at nucleotide positions 821-823 (Figure 75; SEQ ID NO:134). The predicted polypeptide precursor (Figure 76, SEQ ID NO:135) is 228 amino acids long. PRO784 has a calculated molecular weight of approximately 25,735 Daltons and an estimated pI of approximately 5.45. PRO784 has the following features: a signal peptide at about amino acid 1 to about 15; transmembrane domains at about amino acids 68 to about 87 and at about 183 to about 204; potential N-myristoylation sites at about amino acids 15-20, 51-56, 66-60, 163-168, and 206-211; and an RNP-1 protein RNA-binding region at about amino acids 108 to about 117.

Clone DNA53978-1443 was deposited with ATCC on June 16, 1998, and is assigned ATCC deposit no. 209983.

Based on a BLAST and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO784 shows amino acid sequence identity to the following proteins: RNU42209_1,

253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO783 [herein designated as DNA53996-1442] (SEQ ID NO:137) and the derived protein sequence for PRO783.

The entire nucleotide sequence of DNA53996-1442 is shown in Figure 78 (SEQ ID NO:137). Clone
 5 DNA53996-1442 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 310-312 and ending at the stop codon at nucleotide positions 1777-1779 (Figure 78). The predicted polypeptide precursor is 489 amino acids long (Figure 79). The full-length PRO783 protein shown in Figure 79 has an estimated molecular weight of about 55,219 daltons and a pI of about 8.47. Analysis of the full-length PRO783 sequence shown in Figure 79 (SEQ ID NO:138) evidences the presence of the following features:
 10 transmembrane domains located at about amino acids 23-42, 67-89, 111-135, 154-176, 194-218, 296-319, 348-370, 387-410 and 427-452; leucine zipper patterns located at about amino acids 263-283 and 399-420; a potential tyrosine kinase phosphorylation site at about amino acids 180-187; potential N-glycosylation sites at about amino acids 105 -108 and 121-124; potential cAMP- and a cGMP-dependent protein kinase phosphorylation site at about amino acids 288-291; and a region having sequence identity with bacterial
 15 rhodopsins retinal binding site protein at about amino acids 190-218.

An analysis of the Dayhoff database (version 35.45 SwissProt 35) shows some sequence identity between the PRO783 amino acid sequence and the following Dayhoff sequences: YNC2_CAEEL, D64048, ATAC002332_3F4P9.3, NY2R_SHEEP, and VSH_MUMPA.

Clone DNA53996-1442 was deposited with the ATCC on June 2, 1998, and is assigned ATCC deposit
 20 no. 209921.

EXAMPLE 34: Isolation of cDNA Clones Encoding Human PRO820

An expressed sequence tag (EST) DNA database (Merck/Wash. U) was searched and an EST designated EST no. AA504080, Merck clone 825136, was identified (library 312, human B-cell tonsil). Homology searches
 25 revealed that this EST showed sequence identity with low affinity immunoglobulin gamma Fc receptor II. DNA sequencing gave the full-length DNA sequence for PRO820 and the derived protein sequence for PRO820.

The entire nucleotide sequence of DNA56041-1416 is shown in Figure 82 (SEQ ID NO:145). Clone DNA56041-1416 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 115-117 and ending at the stop codon at nucleotide positions 487-489 (Figure 82). The predicted
 30 polypeptide precursor is 124 amino acids long (Figure 83). The full-length PRO820 protein shown in Figure 83 has an estimated molecular weight of about 14,080 daltons and a pI of about 7.48. Clone DNA56041-1416 has been deposited with ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Still analyzing the amino acid sequence of SEQ ID NO:146, the putative signal peptide is at about amino
 35 acids 1-15 of SEQ ID NO:146. Protein kinase C phosphorylation sites are at about amino acids 20-22 and 43-45 of SEQ ID NO:146. An N-myristoylation site is at about amino acids 89-94 of SEQ ID NO:146. An immunoglobulin and major histocompatibility complex domain is at about amino acids 83-90 of SEQ ID NO:146.

The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 35: Isolation of cDNA Clones Encoding Human PRO1080

A consensus DNA sequence was assembled relative to other EST sequences using phrap and was extended using repeated cycles of BLAST and phrap so as to extend the consensus sequence as far as possible using the sources of the EST sequences as described in Example 1 above. The consensus sequence is designated herein as DNA52640. An EST proprietary to Genentech was employed in the consensus assembly and is herein designated as DNA36527 (Figure 86; SEQ ID NO:149).

In light of an observed sequence homology between the DNA36527 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 526423, the Merck EST clone 526423 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 84 and is herein designated as DNA56047-1456.

The entire nucleotide sequence of DNA56047-1456 is shown in Figure 84 (SEQ ID NO:147). Clone DNA56047-1456 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 159-161 and ending at the stop codon at nucleotide positions 1233-1235 of SEQ ID NO:147 (Figure 84). The predicted polypeptide precursor is 358 amino acids long (Figure 85). The full-length PRO1080 protein shown in Figure 85 has an estimated molecular weight of about 40,514 daltons and a pI of about 6.08. Clone DNA56047-1456 has been deposited with ATCC on June 9, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Also shown in Figure 85 are the approximate locations of the signal peptide, cell attachment site, Nt-DnaJ domain signature, region having sequence identity with Nt-DnaJ domain proteins, and N-glycosylation sites. The corresponding nucleic acids of these amino acid sequences and others provided herein can be routinely determined by the information provided herein.

EXAMPLE 36: Isolation of cDNA Clones Encoding Human PRO1079

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above, and is herein designated DNA52714. Based on information provided by the assembly, the clone for Merck EST no. HO6898 was obtained and sequenced, thereby giving the nucleotide sequence designated herein as DNA56050-1455. The entire nucleotide sequence of DNA56050-1455 is shown in Figure 87 (SEQ ID NO:150). Clone DNA56050-1455 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 183-185 and ending at the stop codon at nucleotide positions 861-863 (Figure 87). The predicted polypeptide precursor is 226 amino acids long (Figure 88). The full-length PRO1079 protein shown in Figure 88 has an estimated molecular weight of about 24,611 Daltons and a pI of about 4.85. Analysis of the full-length PRO1079 sequence shown in Figure 88 (SEQ ID NO:3) evidences the presence of the following features: a signal peptide at about amino acid 1-29; potential N-myristoylation sites at about amino acids 10-15, and 51-56; homology to photosystem I psaG and psaK proteins at about amino acids 2 to 20; and homology to prolyl endopeptidase family serine proteins at about amino acids 150 to 163.

Analysis of the amino acid sequence of the full-length PRO1079 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced some sequence identity between the PRO1079 amino acid sequence and the following Dayhoff sequences: CEK10C3_4, MMU50734_1, D69503, AF051149_1, and VSMP_CVMS.

Clone UNQ536 (DNA56050-1455) was deposited with the ATCC on June 22, 1998, and is assigned ATCC deposit no. 203011.

5

EXAMPLE 37: Isolation of cDNA clones Encoding Human PRO793

A cDNA clone (DNA56110-1437) encoding a native human PRO793 polypeptide was identified by a yeast screen, in a human skin tumor cDNA library that preferentially represents the 5' ends of the primary cDNA clones. The yeast screen employed identified a single EST clone designated herein as DNA50177 (Figure 91; SEQ ID NO:154). The DNA50177 sequence was then compared to various EST databases including public EST databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify homologous EST sequences. The comparison was performed using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)]. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). This consensus sequence is herein designated DNA50972.

In light of an observed sequence homology between the DNA50972 consensus sequence and an EST sequence encompassed within the Merck EST clone no. N33874, the Merck EST clone N33874 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 89 and is herein designated as DNA56110-1437.

The full-length DNA56110-1437 clone shown in Figure 89 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 77-79 and ending at the stop codon at nucleotide positions 491-493 (Figure 89). The predicted polypeptide precursor is 138 amino acids long (Figure 90). The full-length PRO793 protein shown in Figure 90 has an estimated molecular weight of about 15,426 daltons and a pI of about 10.67. Analysis of the full-length PRO793 sequence shown in Figure 90 (SEQ ID NO:153) evidences the presence of the following: transmembrane domains from about amino acid 12 to about amino acid 30, from about amino acid 33 to about amino acid 52, from about amino acid 69 to about amino acid 89 and from about amino acid 93 to about amino acid 109, potential N-myristolation sites from about amino acid 11 to about amino acid 16, from about amino acid 51 to about amino acid 56 and from about amino acid 116 to about amino acid 121 and an amino acid sequence block having homology to an aminoacyl-transfer RNA synthetase class-II protein from about amino acid 49 to about amino acid 59. Clone DNA56110-1437 has been deposited with ATCC on August 11, 1998 and is assigned ATCC deposit no. 203113.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 90 (SEQ ID NO:153), evidenced certain homology between the PRO793 amino acid sequence and the following Dayhoff sequences: S47453, AF015193_12, MTEHGNS9_2, E64030, H69784, D64995, CD53_MOUSE, GEN8006, AE001138_7 and COX2_STRPU.

EXAMPLE 38: Isolation of cDNA Clones Encoding Human PRO1016

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. The consensus sequence obtained is herein designated DNA53502.

In light of an observed sequence homology between the DNA53502 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 38680, the Merck EST clone 38680 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 92.

The entire nucleotide sequence of DNA56113-1378 is shown in Figure 92 (SEQ ID NO:155). Clone DNA56113-1378 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 168-170 and ending at the stop codon at nucleotide positions 1302-1304 (Figure 92). The predicted polypeptide precursor is 378 amino acids long (Figure 93). The full-length PRO1016 protein shown in Figure 93 has an estimated molecular weight of about 44,021 daltons and a pI of about 9.07. Clone DNA56113-1378 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1016 polypeptide suggests that portions of it possess sequence identity with acyltransferase, thereby indicating that PRO1016 may be a novel acyltransferase.

Still analyzing the amino acid sequence of SEQ ID NO:156, the putative signal peptide is at about amino acids 1-18 of SEQ ID NO:156. The transmembrane domain(s) are at about amino acids 332-352 and 305-330 of SEQ ID NO:156. The fructose-bisphosphate aldolase class-II protein homology sequence is at about amino acids 73-90 of SEQ ID NO:156. The extradiol ring-cleavage dioxygenase protein is at about amino acids 252-275 of SEQ ID NO:156. The corresponding nucleotides can be routinely determined given the sequences provided herein.

The specific Dayhoff database designation names of sequences to which PRO1016 has sequence identity with include the following: S52645, P_R59712, P_R99249, P_R59713, BNAGPATRF_1, CELT05H4_15 and CELZK40_1.

EXAMPLE 39: Isolation of cDNA Encoding Human PRO1013

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. The consensus DNA sequence was then extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences.

In light of an observed sequence homology between the consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3107695, the Incyte EST clone 3107695 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 94 and is herein designated as DNA56410-1414.

The entire nucleotide sequence of DNA56410-1414 is shown in Figure 94 (SEQ ID NO:157). Clone DNA56410-1414 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 17-19 and ending at the stop codon at nucleotide positions 1244-1246 (Figure 94). The predicted

polypeptide precursor is 409 amino acids long (Figure 95). The full-length PRO1013 protein shown in Figure 95 has an estimated molecular weight of about 46,662 daltons and a pI of about 7.18. Clone DNA56410-1414 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Still analyzing the amino acid sequence of SEQ ID NO:158, the putative signal peptide is at about amino acids 1-19 of SEQ ID NO:158. N-glycosylation sites are at about amino acids 75-78 and 322-325 of SEQ ID NO:158. An N-myristoylation site is at about amino acids 184-189 of SEQ ID NO:158. A growth factor and cytokine receptor family domain is at about amino acids 134-149 of SEQ ID NO:158. The corresponding nucleotides can be routinely determined given the sequences provided herein.

Blast analysis showed some sequence identity with other proteins. Specifically, PRO1013 has some sequence identity with at least the Dayhoff sequences designated: D63877_1; MHU22019_1, AE000730_10, and AF019079_1.

EXAMPLE 40: Isolation of cDNA Clones Encoding Human PRO937

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. That consensus sequence is herein designated DNA49651. Based on the DNA49651 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO937.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CTCCGTGGTAAACCCACAGCCC-3' (SEQ ID NO:161); and
reverse PCR primer 5'-TCACATCGATGGGATCCATGACCG-3' (SEQ ID NO:162).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA48651 sequence which had the following nucleotide sequence:

hybridization probe

5'-GGTCTCGTGACTGTGAAGCCATGTTACAACACTACTGCTCAAACATCATGAG-3' (SEQ ID NO:163).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO937 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO937 [herein designated as DNA56436-1448] (SEQ ID NO:159) and the derived protein sequence for PRO937.

The entire nucleotide sequence of DNA56436-1448 is shown in Figure 96 (SEQ ID NO:159). It contains a single open reading frame having an apparent translational initiation site at nucleotide positions 499-501 and ending at the stop codon found at nucleotide positions 2167-2169 (Figure 96, SEQ ID NO:159). The predicted polypeptide precursor is 556 amino acids long, has a calculated molecular weight of approximately 62,412 daltons and an estimated pI of approximately 6.62. Analysis of the full-length PRO937 sequence shown

in Figure 97 (SEQ ID NO:160) evidences the presence of the following features: signal peptide at about amino acids 1-22; ATP/GTP-binding site motif A (P-loop) at about amino acids 515-523; a potential N-glycosylation site at about amino acids 514-517; and sites of glypican homology at about amino acids 54-74, 106-156, 238-279, 309-345, 423-459, and 468-505.

Clone DNA56436-1448 has been deposited with ATCC on May 27, 1998, and is assigned ATCC deposit no. 209902.

Analysis of the amino acid sequence of the full-length PRO937 polypeptide suggests that it possesses significant sequence similarity to glypican proteins, thereby indicating that PRO937 may be a novel glypican protein. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO937 amino acid sequence and the following Dayhoff sequences: GPCK_MOUSE, GPC2_RAT, GPC5_HUMAN, GPC3_HUMAN, P_R30168, CEC03H12_2, GEN13820, HS119E23_1, HDAC_DROME, and AF017637_1.

EXAMPLE 41: Isolation of cDNA clones Encoding Human PRO842

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence designated herein as Incyte EST cluster sequence no. 69572. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA54230.

In light of an observed sequence homology between the consensus sequence and an EST sequence encompassed within the Merck EST clone no. AA477092, the Merck EST clone AA477092 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 98 and is herein designated as DNA56855-1447.

The full length clone shown in Figure 98 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 153-155 and ending at the stop codon found at nucleotide positions 510-512 (Figure 98; SEQ ID NO:164). The predicted polypeptide precursor (Figure 99, SEQ ID NO:165) is 119 amino acids long. PRO842 has a calculated molecular weight of approximately 13,819 Daltons and an estimated pI of approximately 11.16. Other features of PRO842 include a signal peptide at about amino acids 1-22, a potential protein kinase C phosphorylation site at about amino acids 39-41 and two potential N-myristoylation sites at about amino acids 27-32 and about amino acids 46-51.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 98 (SEQ ID NO:164), evidenced some homology between the PRO842 amino acid sequence and the following Dayhoff sequences: CEZK131_11, P_R80843, RAT5HT2X_1, S81882_1, A60912, MCU60315_137MC137L, U93422_1, p_P91996, U93462_1, and

ZN18_HUMAN.

Clone DNA56855-1447 was deposited with the ATCC on June 23, 1998, and is assigned ATCC deposit no. 203004.

EXAMPLE 42: Isolation of cDNA clones Encoding Human PRO839

5 Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte LIFESEQ® database, designated Incyte EST Cluster No. 24479. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using
10 the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55709.

15 In light of an observed sequence homology between the DNA55709 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 754525, the Merck EST clone 754525 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 100 and is herein designated as DNA56859-1445.

20 The full length clone shown in Figure 100 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 2-4 and ending at the stop codon found at nucleotide positions 263-265 (Figure 100; SEQ ID NO:166). The predicted polypeptide precursor (Figure 101, SEQ ID NO:167) is 87 amino acids long. PRO839 has a calculated molecular weight of approximately 9,719 Daltons and an estimated pI of approximately 4.67. Other features of PRO839 include a signal peptide at about amino acids 1-23, potential protein kinase C phosphorylation sites at about amino acids 37-39 and about amino acids 85-87,
25 a potential casein kinase II phosphorylation site at about amino acids 37-40, sequence identity with ribonucleotide reductase large subunit protein at about amino acids 50-60, and sequence identity with eukaryotic RNA-binding region RNP-1 proteins at about amino acids 70-79.

30 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 101 (SEQ ID NO:167), evidenced some homology between the PRO839 amino acid sequence and the following Dayhoff sequences: CD14_MOUSE, XPR6_YARLI, HS714385_1, S49783, BB19_RABIT, GVPH-HALME, AB003135_1, P_R85453, LUU27081_2, and TP2B_MOUSE.

Clone DNA56859-1445 was deposited with the ATCC on June 23, 1998, and is assigned ATCC deposit no.209019.

EXAMPLE 43: Isolation of cDNA Clones Encoding Human PRO1180

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence (Incyte EST cluster sequence no. 14732). The Incyte EST cluster sequence no. 14732 sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55711.

In light of an observed sequence homology between the DNA55711 consensus sequence and an EST sequence encompassed within the Merck EST clone no. T60981, the Merck EST clone T60981 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 102 and is herein designated DNA56860-1510.

The full length clone shown in Figure 102 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 78-80 and ending at the stop codon found at nucleotide positions 909-911 (Figure 102; SEQ ID NO:168). The predicted polypeptide precursor is 277 amino acids long, has a calculated molecular weight of approximately 31,416 daltons and an estimated pI of approximately 8.88. Analysis of the full-length PRO1180 sequence shown in Figure 103 (SEQ ID NO:169) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 23, a leucine zipper pattern sequence from about amino acid 10 to about amino acid 31, and potential N-myristolation sited from about amino acid 64 to about amino acid 69, from about amino acid 78 to about amino acid 83, from about amino acid 80 to about amino acid 85, from about amino acid 91 to about amino acid 96 and from about amino acid 201 to about amino acid 206. Clone DNA56860-1510 has been deposited with the ATCC on June 9, 1998 and is assigned ATCC deposit no. 209952.

Analysis of the amino acid sequence of the full-length PRO1180 polypeptide suggests that it possesses sequence similarity to the methyltransferase family of proteins. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some degree of homology between the PRO1180 amino acid sequence and the following Dayhoff sequences, MTC165_14, D69267, YH09_YEAST, BIOC_SERMA, ATAC00448415T1D16.16, SHGCPIR_18, SPBC3B9_4, AB009504_14, P_W17977 and A69952.

EXAMPLE 44: Isolation of cDNA clones Encoding Human PRO1134

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 7511. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or

in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55725. Two proprietary Genentech EST sequences were employed in the assembly and are shown in Figure 106 (SEQ ID NO:172) and Figure 107 (SEQ ID NO:173).

5 In light of an observed sequence homology between the DNA55725 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H94897, the Merck EST clone H94897 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 106 and is herein designated as DNA56865-1491.

10 Clone DNA56865-1491 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 153-155 and ending at the stop codon at nucleotide positions 1266-1268 (Figure 104). The predicted polypeptide precursor is 371 amino acids long (Figure 105). The full-length PRO1134 protein shown in Figure 105 has an estimated molecular weight of about 41,935 daltons and a pI of about 9.58. Analysis of the full-length PRO1134 sequence shown in Figure 105 (SEQ ID NO:171) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 23, potential N-glycosylation sites from 15 about amino acid 103 to about amino acid 106, from about amino acid 249 to about amino acid 252 and from about amino acid 257 to about amino acid 260, and an amino acid block having homology to tyrosinase CuA-binding region proteins from about amino acid 280 to about amino acid 306. Clone DNA56865-1491 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203022.

20 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 105 (SEQ ID NO:171), evidenced significant homology between the PRO1134 amino acid sequence and the following Dayhoff sequences: F20P5_18, AC002396_10, S47847, C64146, GSPA_BACSU, P_W10564, RFA1_ECOLI, Y258_HAEIN, RFAJ_SALTY and P_R32985.

25 EXAMPLE 45: Isolation of cDNA clones Encoding Human PRO830

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 20251. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing 30 homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55733.

35 In light of an observed sequence homology between the DNA55733 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H78534, the Merck EST clone H78534 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.

The sequence of this cDNA insert is shown in Figure 108 and is herein designated as DNA56866-1342.

Clone DNA56866-1342 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 154-156 and ending at the stop codon at nucleotide positions 415-417 (Figure 108). The predicted polypeptide precursor is 87 amino acids long (Figure 109). The full-length PRO830 protein shown in Figure 109 has an estimated molecular weight of about 9,272 daltons and a pI of about 9.19. Analysis of the full-length PRO830 sequence shown in Figure 109 (SEQ ID NO:175) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 33, potential N-myristoylation sites from about amino acid 2 to about amino acid 7 and from about amino acid 8 to about amino acid 13 and a thioredoxin family of proteins homology block from about amino acid 23 to about amino acid 39. Clone UNQ470 (DNA56866-1342) has been deposited with ATCC on June 22, 1998 and is assigned ATCC deposit no. 203023.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 109 (SEQ ID NO:175), evidenced significant homology between the PRO830 amino acid sequence and the following Dayhoff sequences: HSU88154_1, HSU88153_1, SAPKSGENE_1, HPU31791_5, GGCNOT2_1, CPU91421_1, CHKESTPC09_1, PQ0769, U91553_79 and B60095.

EXAMPLE 46: Isolation of cDNA clones Encoding Human PRO1115

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the LIFESEQ® database, designated Incyte EST cluster sequence no. 165008. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55726.

In light of an observed sequence homology between the DNA55726 consensus sequence and an EST sequence encompassed within the Merck EST clone no. R75784, the Merck EST clone R75784 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.

The sequence of this cDNA insert is shown in Figure 111 and is herein designated as DNA56868-1478.

The full length clone shown in Figure 110 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 189-191 and ending at the stop codon found at nucleotide positions 1524-1526 (Figure 110; SEQ ID NO:176). The predicted polypeptide precursor (Figure 111, SEQ ID NO:177) is 445 amino acids long. PRO1115 has a calculated molecular weight of approximately 50,533 Daltons and an estimated pI of approximately 8.26. Additional features include a signal peptide at about amino acids 1-20; potential N-glycosylation sites at about amino acids 204-207, 295-298, and 313-316; and putative transmembrane domains at about amino acids 35-54, 75-97, 126-146, 185-204, 333-350, and 353-371.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 111 (SEQ ID NO:177), evidenced some amino acid sequence identity between the PRO1115 amino acid sequence and the following Dayhoff sequences: AF053947_79, S73698, CEC47A10_4, CCOMTND55G_1, HS4LMP2AC_1, LMP2_EBV, PA24_MOUSE, HCU33331_7, P-W05508, and AF002273_1.

5 Clone DNA56868-1478 was deposited with the ATCC on June 23, 1998 and is assigned ATCC deposit no. 203024..

EXAMPLE 47: Isolation of cDNA clones Encoding Human PRO1277

10 A consensus DNA sequence was assembled relative to other ESTs using repeated cycles of BLAST and the program "phrap" as described in Example 1 above. One or more of the ESTs from the assembly was derived from diseased coronary artery tissue. The consensus sequence obtained is designated herein as "DNA49434".

15 In light of an observed sequence homology between the DNA49434 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3042605, the Incyte EST clone 3042605 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 112 (SEQ ID NO:178).

20 Clone DNA56869-1545 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 188-190, and an apparent stop codon at nucleotide positions 2222-2224 (Figure 112). The predicted polypeptide precursor is 678 amino acids long (Figure 113). The full-length PRO1277 protein shown in Figure 113 has an estimated molecular weight of about 73,930 daltons and a pI of about 9.48. Additional features include a signal peptide at about amino acids 1-26; a transmembrane domain at about amino acids 181-200, and potential N-glycosylation sites at about amino acids 390-393 and 520-523.

25 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 113 (SEQ ID NO:179), revealed significant homology between the PRO1277 amino acid sequence and Dayhoff sequence no AF012252_1. Homology was also found between the PRO1277 amino acid sequence and the following Dayhoff sequences: AF006740_1, CA36_HUMAN, HSU1_1, HUMCOL7A1X_1, CA17_HUMAN, MMZ78163_1, CAMA_CHICK, HSU69263_1, YNX3_CAEEL, and MMRNAM3_1.

30 Clone DNA56869-1545 has been deposited with ATCC and is assigned ATCC deposit no. 203161.

EXAMPLE 48: Isolation of cDNA Clones Encoding Human PRO1135

35 A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA52767. Based on the DNA52767 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1135.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with PCR primer pairs prepared based upon the DNA52767 sequence. A positive library was then used to isolate clones encoding the PRO1135 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human coronary artery smooth muscle tissue (LIB309). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1135 [herein designated as DNA56870-1492] (SEQ ID NO:180) and the derived protein sequence for PRO1135.

The entire nucleotide sequence of DNA56870-1492 is shown in Figure 114 (SEQ ID NO:180). Clone DNA56870-1492 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 62-64 and ending at the stop codon at nucleotide positions 1685-1687 (Figure 114). The predicted polypeptide precursor is 541 amino acids long (Figure 115). The full-length PRO1135 protein shown in Figure 115 has an estimated molecular weight of about 60,335 daltons and a pI of about 5.26. Analysis of the full-length PRO1135 sequence shown in Figure 115 (SEQ ID NO:181) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 21, potential N-glycosylation sites from about amino acid 53 to about amino acid 56, from about amino acid 75 to about amino acid 78, from about amino acid 252 to about amino acid 255 and from about amino acid 413 to about amino acid 416 and an amino acid block having homology to glycosyl hydrolase family 35 proteins from about amino acid 399 to about amino acid 414. Clone DNA56870-1492 has been deposited with ATCC on June 2, 1998 and is assigned ATCC deposit no. 209925.

Analysis of the amino acid sequence of the full-length PRO1135 polypeptide suggests that it possesses significant sequence similarity to the alpha 1,2-mannosidase protein, thereby indicating that PRO1135 may be a novel mannosidase. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO1135 amino acid sequence and the following Dayhoff sequences, DMC86E4_5, D86967_1, SPAC23A1_4, YH04_YEAST, B54408, SSMAN9MAN_1, CEZC410_4, S61631 and MSU14190_1.

EXAMPLE 49: Isolation of cDNA Clones Encoding Human PRO1114

A cDNA sequence isolated in the amylase screen described in Example 2 above was found, by the WU-BLAST-2 sequence alignment computer program, to have certain sequence identity to other known interferon receptors. This cDNA sequence is herein designated DNA48466 and is shown in Figure 118 (SEQ ID NO:184). Based on the sequence identity, probes were generated from the sequence of the DNA48466 molecule and used to screen a human breast carcinoma library (LIB135) prepared as described in paragraph 1 of Example 2 above. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes

et al., *Science*, 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp.

The oligonucleotide probes employed were as follows:

forward PCR primer 5'-AGGCTTCGCTGCGACTAGACCTC-3' (SEQ ID NO:185)

reverse PCR primer 5'-CCAGGTCGGGTAAGGATGGTTGAG-3' (SEQ ID NO:186)

hybridization probe

5 5'-TTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC-3' (SEQ ID NO:187)

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 250-252, and a stop signal at nucleotide positions 1183-1185 (Figure 116, SEQ ID NO:182). The predicted polypeptide precursor is 311 amino acids long, has a calculated molecular weight of approximately 35,076 daltons and an estimated pI of approximately 5.04. Analysis of the full-length PRO1114 interferon receptor sequence shown in Figure 117 (SEQ ID NO:183) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 29, a transmembrane domain from about amino acid 230 to about amino acid 255, potential N-glycosylation sites from about amino acid 40 to about amino acid 43 and from about amino acid 134 to about amino acid 137, an amino acid sequence block having homology to tissue factor proteins from about amino acid 92 to about amino acid 119 and an amino acid sequence block having homology to integrin alpha chain proteins from about amino acid 232 to about amino acid 262. Clone DNA57033-1403 has been deposited with ATCC on May 27, 1998 and is assigned ATCC deposit no. 209905.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 117 (SEQ ID NO:183), evidenced significant homology between the PRO1114 interferon receptor amino acid sequence and the following Dayhoff sequences: G01418, INR1_MOUSE, P_R71035, INGS_HUMAN, A26595_1, A26593_1, I56215 and TF_HUMAN.

EXAMPLE 50: Isolation of cDNA Clones Encoding Human PRO828

A consensus DNA sequence was identified using the method described in Example 1 above. This consensus sequence is herein designated DNA35717. Based on the DNA35717 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO828.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GCAGGACTTCTACGACTTCAAGGC-3' (SEQ ID NO:190); and

reverse PCR primer 5'-AGTCTGGGCCAGGTACTTGAAGGC-3' (SEQ ID NO:191).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35717 sequence which had the following nucleotide sequence:

hybridization probe

5'-CAACATCCGGGGCAAACCTGGTGTGCTGGAGAAGTACCGCGGATCGGTGT-3' (SEQ ID NO:192)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO828 gene using the probe oligonucleotide and one of the PCR primers. RNA

for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB25).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO828 [herein designated as DNA57037-1444] (SEQ ID NO:188) and the derived protein sequence for PRO828.

The entire nucleotide sequence of DNA57037-1444 is shown in Figure 119 (SEQ ID NO:188). Clone DNA57037-1444 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 34-36 and ending at the stop codon at nucleotide positions 595-597 (Figure 119). The predicted polypeptide precursor is 187 amino acids long (Figure 120). The full-length PRO828 protein shown in Figure 120 has an estimated molecular weight of about 20,996 daltons and a pI of about 8.62. Analysis of the full-length PRO828 sequence shown in Figure 120 (SEQ ID NO:189) evidences the presence of the following: a signal peptide at about amino acids 1- 21; sequences identity to glutathione peroxidases signature 2 at about amino acids 82-89; sequence identity to glutathione peroxidases selenocysteine proteins at about amino acids 35-60, 63-100, 107-134, and 138-159. Clone DNA57037-1444 has been deposited with ATCC on May 27, 1998, and is assigned ATCC deposit no. 209903.

Analysis of the amino acid sequence of the full-length PRO828 polypeptide suggests that it possesses significant sequence similarity to glutathione peroxidases, thereby indicating that PRO828 may be a novel peroxidase enzyme. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced sequence identity between the PRO828 amino acid sequence and the following Dayhoff sequences: AF053311_1, CELT09A12_2, AC004151_3, BTUE_ECOLI, CER05H10_3, P_P80918, PWU88907_1, and P_W22308.

EXAMPLE 51: Isolation of cDNA clones Encoding Human PRO1009

A cDNA clone (DNA57129-1413) encoding a native human PRO1009 polypeptide was identified by the use of a yeast screen, in a human SK-Lu-1 adenocarcinoma cell line cDNA library that preferentially represents the 5' ends of the primary cDNA clones. First SEQ ID NO:195 (Figure 123) was identified, which was extended by alignments to other EST sequences to form a consensus sequence. Oligonucleotide probes based upon the consensus sequence were synthesized and used to screen the cDNA library which gave rise to the full-length DNA57129-1413 clone.

The full length DNA57129-1413 clone shown in Figure 121 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 41-43 and ending at the stop codon found at nucleotide positions 1886-1888 (Figure 121; SEQ ID NO:193). The predicted polypeptide precursor (Figure 122, SEQ ID NO:194) is 615 amino acids long. Figure 122 also shows the approximate locations of the signal sequence, transmembrane domains, myristoylation sites, a glycosylation site and an AMP-binding domain. PRO1009 has a calculated molecular weight of approximately 68,125 daltons and an estimated pI of approximately 7.82. Clone DNA57129-1413 has been deposited with ATCC and is assigned ATCC deposit no. 209977. It is understood that the deposited clone has the actual and correct sequence and that the representations herein may have minor, normal sequencing errors.

Based on a WU-BLAST-2 sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO1009 shows amino acid sequence identity to at least the following proteins which were designated in a Dayhoff database as follows: F69893, CEF28F8_2, BSY13917_7, BSY13917_7, D69187, D69649, XCRPFB_1, E64928, YDID_ECOLI, BNACSF8_1 and RPU75363_2.

5 EXAMPLE 52: Isolation of cDNA Clones Encoding Human PRO1007

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated as DNA40671.

10 In light of an observed sequence homology between the DNA40671 consensus sequence and an EST sequence encompassed within the Merck EST clone no. T70513, the Merck EST clone T70513 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 124.

The entire nucleotide sequence of DNA57690-1374 is shown in Figure 124 (SEQ ID NO:196). Clone DNA57690-1374 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 16-18 and ending at the stop codon at nucleotide positions 1054-1056 (Figure 124). The predicted
15 polypeptide precursor is 346 amino acids long (Figure 125). The full-length PRO1007 protein shown in Figure 125 has an estimated molecular weight of about 35,971 daltons and a pI of about 8.17. Clone DNA57690-1374 has been deposited with the ATCC on June 9, 1998. Regarding the sequence, it is understood that the deposited clone contains the actual sequence, and the sequences provided herein are based on known sequencing techniques. The representative figures herein show the representative numbering.

20 Analysis of the amino acid sequence of the full-length PRO1007 polypeptide suggests that portions of it possess sequence identity to MAGPIAP, thereby indicating that PRO1007 may be a novel member of the family to which MAGPIAP belongs.

Still analyzing the amino acid sequence of SEQ ID NO:197, the putative signal peptide is at about amino acids 1-30 of SEQ ID NO:197. The transmembrane domain is at amino acids 325-346 of SEQ ID NO:197. N-glycosylation sites are at about amino acids 118-121, 129-132, 163-166, 176-179, 183-186 and 227-130 of SEQ
25 ID NO:197. Ly-6/u-Par domain protein homology is at about amino acids 17-36 and 209-222 of SEQ ID NO:197. The corresponding nucleotides of the amino acids presented herein can be routinely determined given the sequences provided herein.

30 EXAMPLE 53: Isolation of cDNA clones Encoding Human PRO1056

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated herein as 6425. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify
35 existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a

consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55736.

In light of an observed sequence homology between the DNA55736 consensus sequence and an EST sequence encompassed within the Merck EST clone no. R88049, the Merck EST clone R88049 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.

5 The sequence of this cDNA insert is shown in Figure 126 and is herein designated as DNA57693-1424.

Clone DNA57693-1424 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 56-58 and ending at the stop codon at nucleotide positions 416-418 (Figure 126). The predicted polypeptide precursor is 120 amino acids long (Figure 127). The full-length PRO1056 protein shown in Figure 127 has an estimated molecular weight of about 13,345 daltons and a pI of about 5.18. Analysis of
10 the full-length PRO1056 sequence shown in Figure 127 (SEQ ID NO:199) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 18, a transmembrane domain from about amino acid 39 to about amino acid 58, a potential N-glycosylation site from about amino acid 86 to about amino acid 89, protein kinase C phosphorylation sites from about amino acid 36 to about amino acid 38 and from about amino acid 58 to about amino acid 60, a tyrosine kinase phosphorylation site from about amino acid 25 to about
15 amino acid 32 and an amino acid sequence block having homology to channel forming colicin proteins from about amino acid 24 to about amino acid 56. Clone DNA57693-1424 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203008.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 127 (SEQ ID NO:199), evidenced significant
20 homology between the PRO1056 amino acid sequence and the following Dayhoff sequences: PLM_HUMAN, A40533, ATNG_HUMAN, A55571, ATNG_SHEEP, S31524, GEN13025, RIC_MOUSE, A48678 and A10871_1.

EXAMPLE 54: Isolation of cDNA clones Encoding Human PRO826

25 Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 47283. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul
30 et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56000.

In light of an observed sequence homology between the DNA56000 consensus sequence and an EST
35 sequence encompassed within the Merck EST clone no. W69233, the Merck EST clone W69233 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 128 and is herein designated as DNA57694-1341.

Clone DNA57694-1341 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 13-15 and ending at the stop codon at nucleotide positions 310-312 (Figure 128). The predicted polypeptide precursor is 99 amino acids long (Figure 129). The full-length PRO826 protein shown in Figure 129 has an estimated molecular weight of about 11,050 daltons and a pI of about 7.47. Analysis of the full-length PRO826 sequence shown in Figure 129 (SEQ ID NO:201) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 22, potential N-myristoylation sites from about amino acid 22 to about amino acid 27 and from about amino acid 90 to about amino acid 95 and an amino acid sequence block having homology to peroxidase from about amino acid 16 to about amino acid 48. Clone DNA57694-1341 has been deposited with ATCC on June 22, 1998 and is assigned ATCC deposit no. 203017.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 129 (SEQ ID NO:201), evidenced significant homology between the PRO826 amino acid sequence and the following Dayhoff sequences: CCU12315_1, SCU96108_6, CELF39F10_4 and HELT_HELHO.

EXAMPLE 55: Isolation of cDNA clones Encoding Human PRO819

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 49605. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56015.

In light of an observed sequence homology between the DNA56015 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H65785, the Merck EST clone H65785 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 130 and is herein designated as DNA57695-1340.

Clone DNA57695-1340 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 46-48 and ending at the stop codon at nucleotide positions 202-204 (Figure 130). The predicted polypeptide precursor is 52 amino acids long (Figure 131). The full-length PRO819 protein shown in Figure 131 has an estimated molecular weight of about 5,216 daltons and a pI of about 4.67. Analysis of the full-length PRO819 sequence shown in Figure 131 (SEQ ID NO:203) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 24, a potential N-myristoylation site from about amino acid 2 to about amino acid 7 and a region having homology to immunoglobulin light chain from about amino acid 5 to about amino acid 33. Clone DNA57695-1340 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203006.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 131 (SEQ ID NO:203), evidenced significant homology between the PRO819 amino acid sequence and the following Dayhoff sequences: HSU03899_1, HUMIGLITEB_1, VG28_HSVSA, AF031522_1, PAD1_YEAST and AF045484_1.

5 EXAMPLE 56: Isolation of cDNA Clones Encoding Human PRO1006

An initial candidate sequence from Incyte cluster sequence no. 45748 was identified using the signal algorithm process described in Example 3 above. This sequence was then aligned with a variety of public and Incyte EST sequences and a consensus sequence designated herein as DNA56036 was derived therefrom.

10 In light of an observed sequence homology between the DNA56036 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 489737, the Merck EST clone 489737 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 132.

15 The entire nucleotide sequence of DNA57699-1412 is shown in Figure 132 (SEQ ID NO:204). Clone DNA57699-1412 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 28-30 and ending at the stop codon at nucleotide positions 1204-1206 (Figure 132). The predicted polypeptide precursor is 392 amino acids long (Figure 133). The full-length PRO1006 protein shown in Figure 133 has an estimated molecular weight of about 46,189 daltons and a pI of about 9.04. Clone DNA57699-1412 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

20 Analyzing the amino acid sequence of SEQ ID NO:205, the putative signal peptide is at about amino acids 1-23 of SEQ ID NO:205. The N-glycosylation sites are at about amino acids 40-43, 53-56, 204-207 and 373-376 of SEQ ID NO:205. An N-myristoylation site is at about amino acids 273-278 of SEQ ID NO:205.

The corresponding nucleotides of these amino acid regions and others can be routinely determined given the sequences provided herein.

25

EXAMPLE 57: Isolation of cDNA Clones Encoding Human PRO1112

Use of the signal sequence algorithm described in Example 3 above allowed identification of a specific EST cluster sequence. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56018.

35

In light of an observed sequence homology between the DNA56018 consensus sequence and an EST sequence encompassed within the Merck EST clone no. AA223546, the Merck EST clone AA223546 was

purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 134 and is herein designated as DNA57702-1476.

The entire nucleotide sequence of DNA57702-1476 is shown in Figure 134 (SEQ ID NO:206). Clone DNA57702-1476 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 20-22 and ending at the stop codon at nucleotide positions 806-808 of SEQ ID NO:206 (Figure 134).

The predicted polypeptide precursor is 262 amino acids long (Figure 135). The full-length PRO1112 protein shown in Figure 135 has an estimated molecular weight of about 29,379 daltons and a pI of about 8.93. Figure 135 also shows the approximate locations of the signal peptide and transmembrane domains. Clone DNA57702-1476 has been deposited with the ATCC on June 9, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1112 polypeptide suggests that it possesses some sequence similarity to other proteins. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some sequence identity between the PRO1112 amino acid sequence and at least the following Dayhoff sequences, MTY20B11_13 (a mycobacterium tuberculosis peptide), F64471, AE000690_6, XLU16364_1, E43259 (H⁺-transporting ATP synthase) and PIGSLADRXE_1 (MHC class II histocompatibility antigen).

EXAMPLE 58: Isolation of cDNA clones Encoding Human PRO1074

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence (Incyte cluster sequence No. 42586). This cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQTM, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, Univ. of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56251.

In light of an observed sequence homology between the DNA56251 consensus sequence and an EST sequence encompassed within the Merck EST clone no. AA081912, the Merck EST clone AA081912 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 136 and is the full-length DNA sequence for PRO1074. Clone DNA57704-1452 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209953.

The entire nucleotide sequence of DNA57704-1452 is shown in Figure 136 (SEQ ID NO:208). Clone DNA57704-1452 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 322-324 and ending at the stop codon at nucleotide positions 1315-1317 (Figure 136). The predicted polypeptide precursor is 331 amino acids long (Figure 137). The full-length PRO1074 protein shown in Figure 137 has an estimated molecular weight of about 39,512 Daltons and a pI of about 8.03. Analysis of the full-

length PRO1074 sequence shown in Figure 137 (SEQ ID NO:209) evidences the presence of the following features: a transmembrane domain at about amino acids 20 to 39; potential N-glycosylation sites at about amino acids 72 to 75, 154 to 157, 198 to 201, 212 to 215, and 326 to 329; a glycosaminoglycan attachment site at about amino acids 239 to 242, and a Ly-6/u-PAR domain at about amino acids 23 to 36.

Analysis of the amino acid sequence of the full-length PRO1074 polypeptide suggests that it possesses significant sequence similarity to beta 1,3-galactosyltransferase, thereby indicating that PRO1074 may be a novel member of the galactosyltransferase family of proteins. Analysis of the amino acid sequence of the full-length PRO1074 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1074 amino acid sequence and the following Dayhoff sequences: AF029792_1, P_R57433, DMU41449_1, AC000348_14, P_R47479, CET09F5_2, CEF14B6_4, CET15D6_5, CEC54C8_4, and CEE03H4_10.

Clone DNA57704-1452 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209953.

EXAMPLE 59: Isolation of cDNA clones Encoding Human PRO1005

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the LIFESEQ® database, Incyte cluster sequence no. 49243. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56380.

In light of an observed sequence homology between the DNA56380 consensus sequence and an EST sequence encompassed within the Merck EST clone no. AA256657, the Merck EST clone AA256657 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 138 and is herein designated as DNA57708-1411.

The full length clone shown in Figure 138 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 30-32 and ending at the stop codon found at nucleotide positions 585-587 (Figure 138; SEQ ID NO:210). The predicted polypeptide precursor (Figure 139, SEQ ID NO:211) is 185 amino acids long. PRO1005 has a calculated molecular weight of approximately 20,331 daltons and an estimated pI of approximately 5.85. Clone DNA57708-1411 was deposited with the ATCC June 23, 1998, and is assigned ATCC deposit no. 203021.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 139 (SEQ ID NO:211), evidenced some homology between the PRO1005 amino acid sequence and the following Dayhoff sequences: DDU07187_1, DDU87912_1, CELD1007_14, A42239, DDU42597_1, CYAG_DICDI, S50452, MRKC_KLEPN, P-R41998,

and XYNA_RUMFL.

EXAMPLE 60: Isolation of cDNA clones Encoding Human PRO1073

An initial DNA sequence referred to herein as DNA55938 and shown in Figure 142 (SEQ ID NO:214) was identified using a yeast screen, in a human SK-Lu-1 adenocarcinoma cell line cDNA library that preferentially represents the 5' ends of the primary cDNA clones. DNA55938 was then compared to ESTs from public databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA), using the computer program BLAST or BLAST2 [Altschul et al., *Methods in Enzymology*, 266:460-480 (1996)]. The ESTs were clustered and assembled into a consensus DNA sequence using the computer program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained is designated herein as DNA56411.

In light of an observed sequence homology between the DNA56411 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H86027, the Merck EST clone H86027 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 140.

The full length DNA57710-1451 clone shown in Figure 140 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 345-347 and ending at the stop codon found at nucleotide positions 1242-1244 (Figure 140; SEQ ID NO:212). The predicted polypeptide precursor (Figure 141, SEQ ID NO:213) is 299 amino acids long. PRO1073 has a calculated molecular weight of approximately 34,689 daltons and an estimated pI of approximately 11.49. The PRO1073 polypeptide has the following additional features: a signal peptide at about amino acids 1-31, sequence identity to bZIP transcription factor basic domain signature at about amino acids, a potential N-glycosylation site at about amino acids 2-5, and sequence identity with protamine P1 proteins at about amino acids 158-183.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 141 (SEQ ID NO:213), revealed some sequence identity between the PRO1073 amino acid sequence and the following Dayhoff sequences: MMU37351_1, ATAC00250510T9J22.10, S59043, ENXNUPR_1, B47328, SR55_DROME, S26650, SON_HUMAN, VIT2_CHICK, and XLC4SRPRT_1.

Clone DNA57710-1451 was deposited with the ATCC on July 1, 1998 and is assigned ATCC deposit no. 203048.

EXAMPLE 61: Isolation of cDNA clones Encoding Human PRO1152

A cDNA clone (DNA57711-1501) encoding a native human PRO1152 polypeptide was identified by employing a yeast screen, in a human infant brain cDNA library that preferentially represents the 5' ends of the primary cDNA clones. Specifically, a yeast screen was employed to identify a cDNA designated herein as DNA55807 (SEQ ID NO:217; see Figure 145).

In light of an observed sequence homology between the DNA55807 sequence and an EST sequence encompassed within the Merck EST clone no. R56756, the Merck EST clone R56756 was purchased and the

cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 143.

The full-length DNA57711-1501 clone shown in Figure 143 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 58-60 and ending at the stop codon at nucleotide positions 1495-1497 (Figure 143). The predicted polypeptide precursor is 479 amino acids long (Figure 144).

5 The full-length PRO1152 protein shown in Figure 144 has an estimated molecular weight of about 53,602 daltons and a pI of about 8.82. Analysis of the full-length PRO1152 sequence shown in Figure 144 (SEQ ID NO:216) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 28, transmembrane domains from about amino acid 133 to about amino acid 155, from about amino acid 168 to about amino acid 187, from about amino acid 229 to about amino acid 247, from about amino acid 264 to about amino acid 285, from about amino acid 309 to about amino acid 330, from about amino acid 371 to about amino acid 390 and from about amino acid 441 to about amino acid 464, potential N-glycosylation sites from about amino acid 34 to about amino acid 37 and from about amino acid 387 to about amino acid 390 and an amino acid sequence block having homology to a respiratory-chain NADH dehydrogenase subunit from about amino acid 245 to about amino acid 287. Clone DNA57711-1501 has been deposited with ATCC on July 1, 1998 and is assigned ATCC deposit no. 203047.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 144 (SEQ ID NO:216), evidenced significant homology between the PRO1152 amino acid sequence and the following Dayhoff sequences: AF052239_1, SYNN9CGA_1, SFCYTB2_1, GEN12507, P_R11769, MTV025_109, C61168, S43171, P_P61689 and P_P61696.

EXAMPLE 62: Isolation of cDNA clones Encoding Human PRO1136

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 109142. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56039.

In light of an observed sequence homology between the DNA56039 consensus sequence and an EST sequence encompassed within the Merck EST clone no. HSC1NF011, the Merck EST clone HSC1NF011 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 146 and is herein designated as DNA57827-1493.

Clone DNA57827-1493 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 216-218 and ending at the stop codon at nucleotide positions 2112-2114 (Figure 146).

The predicted polypeptide precursor is 632 amino acids long (Figure 147). The full-length PRO1136 protein shown in Figure 147 has an estimated molecular weight of about 69,643 daltons and a pI of about 8.5. Analysis of the full-length PRO1136 sequence shown in Figure 147 (SEQ ID NO:219) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 15 and potential N-glycosylation sites from about amino acid 108 to about amino acid 111, from about amino acid 157 to about amino acid 160, from about amino acid 289 to about amino acid 292 and from about amino acid 384 to about amino acid 387. Clone DNA57827-1493 has been deposited with ATCC on July 1, 1998 and is assigned ATCC deposit no. 203045.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 147 (SEQ ID NO:219), evidenced significant homology between the PRO1136 amino acid sequence and the following Dayhoff sequences: AF034746_1, AF034745_1, MMAF000168_19, HSMUPP1_1, AF060539_1, SP97_RAT, I38757, MMU93309_1, CEK01A6_4 and HSA224747_1.

EXAMPLE 63: Isolation of cDNA clones Encoding Human PRO813

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence (Incyte EST cluster sequence no. 45501. The Incyte EST cluster sequence no. 45501 sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56400.

In light of an observed sequence homology between the DNA56400 consensus sequence and an EST sequence encompassed within the Merck EST clone no. T90592, the Merck EST clone T90592 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 148 and is herein designated DNA57834-1339.

The full length clone shown in Figure 148 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 109-111 and ending at the stop codon found at nucleotide positions 637-639 (Figure 149; SEQ ID NO:221). The predicted polypeptide precursor is 176 amino acids long, has a calculated molecular weight of approximately 19,616 daltons and an estimated pI of approximately 7.11. Analysis of the full-length PRO813 sequence shown in Figure 149 (SEQ ID NO:221) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 26 and potential N-myristoylation sites from about amino acid 48 to about amino acid 53, from about amino acid 153 to about amino acid 158, from about amino acid 156 to about amino acid 161 and from about amino acid 167 to about amino acid 172. Clone DNA57834-1339 has been deposited with the ATCC on June 9, 1998 and is assigned ATCC deposit no. 209954.

Analysis of the amino acid sequence of the full-length PRO813 polypeptide suggests that it possesses sequence similarity to the pulmonary surfactant-associated protein C. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some degree of homology between the PRO813 amino acid sequence and the following Dayhoff sequences, PSPC_MUSVI, P_P92071, G02964, P_R65489, P_P82977, P_R84555, S55542, MUSIGHAJ_1 and PH1158.

5

EXAMPLE 64: Isolation of cDNA Clones Encoding Human PRO809

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence. The Incyte EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56418.

15

In light of an observed sequence homology between the DNA56418 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H74302, the Merck EST clone H74302 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 150 and is herein designated DNA57836-1338.

20

The entire nucleotide sequence of DNA57836-1338 is shown in Figure 150 (SEQ ID NO:222). Clone DNA57836-1338 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 63-65 and ending at the stop codon at nucleotide positions 858-860 of SEQ ID NO:222 (Figure 150). The predicted polypeptide precursor is 265 amino acids long (Figure 151). The full-length PRO809 protein shown in Figure 151 has an estimated molecular weight of about 29,061 daltons and a pI of about 9.18. Figure 151 further shows the approximate positions of the signal peptide and N-glycosylation sites. The corresponding nucleotides can be determined by referencing Figure 150. Clone DNA57836-1338 has been deposited with ATCC on June 23, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

25

Analysis of the amino acid sequence of the full-length PRO809 polypeptide suggests that it possesses some sequence similarity to the heparin sulfate proteoglycan and to endothelial cell adhesion molecule-1. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced sequence identity between the PRO809 amino acid sequence and the following Dayhoff sequences, PGBM_MOUSE, D82082_1 and PW14158.

30

35 EXAMPLE 65: Isolation of cDNA Clones Encoding Human PRO791

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence. The Incyte EST cluster sequence was then compared to a variety of expressed

sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56429.

In light of an observed sequence homology between the DNA56429 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 36367, the Merck EST clone 36367 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 152 and is herein designated DNA57838-1337.

The entire nucleotide sequence of DNA57838-1337 is shown in Figure 152 (SEQ ID NO:224). Clone DNA57838-1337 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 9-11 and ending at the stop codon at nucleotide positions 747-749 of SEQ ID NO:224 (Figure 152). The predicted polypeptide precursor is 246 amino acids long (Figure 153). The full-length PRO791 protein shown in Figure 153 has an estimated molecular weight of about 27,368 daltons and a pI of about 7.45. Figure 153 also shows the approximate locations of the signal peptide, the transmembrane domain, N-glycosylation sites and a region conserved in extracellular proteins. The corresponding nucleotides of one embodiment provided herein can be identified by referencing Figure 152. Clone DNA57838-1337 has been deposited with ATCC on June 23, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO791 polypeptide suggests that it has sequence similarity with MHC-I antigens, thereby indicating that PRO791 may be related to MHC-I antigens. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some sequence identity between the PRO791 amino acid sequence and the following Dayhoff sequences, AF034346_1, MMQ1K5_1 and HFE_HUMAN.

EXAMPLE 66: Isolation of cDNA clones Encoding Human PRO1004

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence, Incyte cluster sequence No. 73681. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, Univ. of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated as DNA56516.

In light of an observed sequence homology between the DNA56516 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H43837, the Merck EST clone H43837 was purchased

and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 154.

The full length clone shown in Figure 154 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 119-121 and ending at the stop codon at nucleotide positions 464-466 (Figure 154; SEQ ID NO:226). The predicted polypeptide precursor is 115 amino acids long (Figure 155; SEQ ID NO:227). The full-length PRO1004 protein shown in Figure 155 has an estimated molecular weight of about 13,649 daltons and a pI of about 9.58. Analysis of the full-length PRO1004 sequence shown in Figure 155 (SEQ ID NO:227) evidences the presence of the following features: a signal peptide at about amino acids 1-24, a microbodies C-terminal targeting signal at about amino acids 113-115, a potential N-glycosylation site at about amino acids 71-74, and a domain having sequence identity with dihydrofolate reductase proteins at about amino acids 22-48.

Analysis of the amino acid sequence of the full-length PRO1004 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1004 amino acid sequence and the following Dayhoff sequences: CELR02D3_7, LEC1_MOUSE, AF006691_3, SSZ97390_1, SSZ97395_1, and SSZ97400_1.

Clone DNA57844-1410 was deposited with the ATCC on June 23, 1998, and is assigned ATCC deposit no. 203010.

EXAMPLE 67: Isolation of cDNA clones Encoding Human PRO1111

An expressed sequence tag (EST) DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST was identified which had homology to insulin-like growth factor binding protein.

RNA for construction of cDNA libraries was isolated from human fetal brain. The cDNA libraries used to isolate the cDNA clones encoding human PRO1111 were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI.

The human fetal brain cDNA libraries (prepared as described above), were screened by hybridization with a synthetic oligonucleotide probe based upon the Incyte EST sequence described above:

5'-CCACCACCTGGAGGTCCTGCAGTTGGGCAGGAAGTCCATCCGGCAGATTG-3' (SEQ ID NO:251).

An identified cDNA clone was sequenced in entirety. The entire nucleotide sequence of PRO1111 is shown in Figure 156 (SEQ ID NO:228). Clone DNA58721-1475 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 57-59 and a stop codon at nucleotide positions 2016-2018 (Figure 156; SEQ ID NO:228). The predicted polypeptide precursor is 653 amino acids long (Figure 157). The transmembrane domains are at positions 21-40 (type II) and 528-548. Clone DNA58721-1475 has been deposited with ATCC and is assigned ATCC deposit no. 203110. The full-length PRO1111 protein shown in Figure 157 has an estimated molecular weight of about 72,717 daltons and a pI of about 6.99.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 157 (SEQ ID NO:229), revealed some sequence identity between the PRO1111 amino acid sequence and the following Dayhoff sequences: A58532, D86983_1, RNPLGPV_1, PGS2_HUMAN, AF038127_1, ALS_MOUSE, GPV_HUMAN, PGS2_BOVIN, ALS_PAPPA and I47020.

5

EXAMPLE 68: Isolation of cDNA clones Encoding Human PRO1344

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA33790. Based on the DNA33790 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1344.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AGGTTTCGTGATGGAGACAACCGCG-3' (SEQ ID NO:232)

reverse PCR primer 5'-TGTC AAGGACGCACTGCCGTCATG-3' (SEQ ID NO:233)

15 Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA33790 sequence which had the following nucleotide sequence

hybridization probe

5'-TGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCC-3' (SEQ ID NO:234)

20 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1344 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

25 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1344 (designated herein as DNA58723-1588 [Figure 158, SEQ ID NO:230]); and the derived protein sequence for PRO1344.

30 The entire nucleotide sequence of DNA58723-1588 is shown in Figure 158 (SEQ ID NO:230). Clone DNA58723-1588 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 26-28 and ending at the stop codon at nucleotide positions 2186-2188 (Figure 158). The predicted polypeptide precursor is 720 amino acids long (Figure 159). The full-length PRO1344 protein shown in Figure 159 has an estimated molecular weight of about 80,199 daltons and a pI of about 7.77. Analysis of the full-length PRO1344 sequence shown in Figure 159 (SEQ ID NO:231) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 23, an EGF-like domain cysteine protein signature sequence from about amino acid 260 to about amino acid 271, potential N-glycosylation sites from about amino acid 96 to about amino acid 99, from about amino acid 279 to about amino acid 282, from about amino acid 316 to about amino acid 319, from about amino acid 451 to about amino acid 454 and from about amino acid 614 to about amino acid 617, an amino acid sequence block having homology to serine proteases, trypsin family from about amino acid 489 to about amino acid 505 and a CUB domain protein profile sequence from about amino

acid 150 to about amino acid 166. Clone DNA58723-1588 has been deposited with ATCC on August 18, 1998 and is assigned ATCC deposit no. 203133.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 159 (SEQ ID NO:231), evidenced significant homology between the PRO1344 amino acid sequence and the following Dayhoff sequences: S77063_1, CRAR_MOUSE, P_R74775, P_P90070, P_R09217, P_P70475, HSBMP16_1 and U50330_1.

EXAMPLE 69: Isolation of cDNA clones Encoding Human PRO1109

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA52642. The consensus DNA sequence was obtained by extending using repeated cycles of BLAST and phrap a previously obtained consensus sequence as far as possible using the sources of EST sequences discussed above. Based on the DNA52642 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1109.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCTTACCTCAGAGGCCAGAGCAAGC-3' (SEQ ID NO:237)

reverse PCR primer 5'-GAGCTTCATCCGTTCTGCGTTCACC-3' (SEQ ID NO:238)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA52642 sequence which had the following nucleotide sequence

hybridization probe

5'-CAGGAATGTAAAGCTTTACAGAGGGTCGCCATCCTCGTTCCCCACC-3' (SEQ ID NO:239)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1109 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human SK-Lu-1 adenocarcinoma cell tissue (LIB247).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1109 (designated herein as DNA58737-1473 [Figure 160, SEQ ID NO:235]) and the derived protein sequence for PRO1109.

The entire nucleotide sequence of DNA58737-1473 is shown in Figure 160 (SEQ ID NO:235). Clone DNA58737-1473 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 119-120 and ending at the stop codon at nucleotide positions 1151-1153 (Figure 160). The predicted polypeptide precursor is 344 amino acids long (Figure 161). The full-length PRO1109 protein shown in Figure 161 has an estimated molecular weight of about 40,041 daltons and a pI of about 9.34. Analysis of the full-length PRO1109 sequence shown in Figure 161 (SEQ ID NO:236) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 27, potential N-glycosylation sites from about amino acid 4 to about amino acid 7, from about amino acid 220 to about amino acid 223 and from about amino acid 335 to about amino acid 338 and an amino acid sequence block having homology to xylose isomerase proteins from about amino acid 191 to about amino acid 201. Clone DNA58737-1473 has been deposited with ATCC

on August 18, 1998 and is assigned ATCC deposit no. 203136.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 161 (SEQ ID NO:236), evidenced significant homology between the PRO1109 amino acid sequence and the following Dayhoff sequences: HSUDPGAL_1, HSUDP14_1, NALS_BOVIN, HSU10473_1, CEW02B12_11, YNJ4_CAEEL, AE000738_11, CET24D1_1, S48121 and CEGLY9_1.

EXAMPLE 70: Isolation of cDNA clones Encoding Human PRO1383

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA53961. Based on the DNA53961 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1383.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CATTTCTTACCCTGGACCCAGCTCC-3' (SEQ ID NO:242)

reverse PCR primer 5'-GAAAGGCCCCACAGCACATCTGGCAG-3' (SEQ ID NO:243)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA53961 sequence which had the following nucleotide sequence

hybridization probe

5'-CCACGACCCGAGCAACTTCCTCAAGACCGACTTGTTTCTCTACAGC-3' (SEQ ID NO:244)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1383 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1383 (designated herein as DNA58743-1609 [Figure 162, SEQ ID NO: 240]) and the derived protein sequence for PRO1383.

The entire nucleotide sequence of DNA58743-1609 is shown in Figure 162 (SEQ ID NO:240). Clone DNA58743-1609 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 122-124 and ending at the stop codon at nucleotide positions 1391-1393 (Figure 162). The predicted polypeptide precursor is 423 amino acids long (Figure 163). The full-length PRO1383 protein shown in Figure 163 has an estimated molecular weight of about 46,989 daltons and a pI of about 6.77. Analysis of the full-length PRO1383 sequence shown in Figure 163 (SEQ ID NO:241) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 24, a transmembrane domain from about amino acid 339 to about amino acid 362, and potential N-glycosylation sites from about amino acid 34 to about amino acid 37, from about amino acid 58 to about amino acid 61, from about amino acid 142 to about amino acid 145, from about amino acid 197 to about amino acid 200, from about amino acid 300 to about amino acid 303 and from about amino acid 364 to about amino acid 367. Clone DNA58743-1609 has been deposited with ATCC on

August 25, 1998 and is assigned ATCC deposit no. 203154.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 163 (SEQ ID NO:241), evidenced significant homology between the PRO1383 amino acid sequence and the following Dayhoff sequences: NMB_HUMAN, QNR_COTJA, P_W38335, P115_CHICK, P_W38164, A45993_1, MMU70209_1, D83704_1 and P_W39176.

5

EXAMPLE 71: Isolation of cDNA Clones Encoding Human PRO1003

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence designated herein as 43055. This sequence was then compared to a variety of EST databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated consen01.

15

In light of an observed sequence homology between the consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2849382, the Incyte EST clone 2849382 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 164.

20

The entire nucleotide sequence of DNA58846-1409 is shown in Figure 164 (SEQ ID NO:245). Clone DNA58846-1409 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 41-43 and ending at the stop codon at nucleotide positions 293-295 (Figure 164). The predicted polypeptide precursor is 84 amino acids long (Figure 165). The full-length PRO1003 protein shown in Figure 165 has an estimated molecular weight of about 9,408 daltons and a pI of about 9.28. Analysis of the full-length PRO1003 sequence shown in Figure 165 (SEQ ID NO:246) evidences the presence of a signal peptide at amino acids 1 to about 24, and a cAMP- and cGMP-dependent protein kinase phosphorylation site at about amino acids 58 to about 61. Analysis of the amino acid sequence of the full-length PRO1003 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1003 amino acid sequence and the following Dayhoff sequences: AOPCZA363_3, SRTX_ATREN, A48298, MHVJHMS_1, VGL2_CVMJH, DHDHTC2_2, CORT_RAT, TAL6_HUMAN, P_W14123, and DVUFI_2.

25

30

EXAMPLE 72: Isolation of cDNA Clones Encoding Human PRO1108

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA53237.

35

In light of an observed sequence homology between the DNA53237 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2379881, the Incyte EST clone 2379881 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.

The sequence of this cDNA insert is shown in Figure 166 and is herein designated DNA58848-1472.

The entire nucleotide sequence of DNA58848-1472 is shown in Figure 166 (SEQ ID NO:247). Clone DNA58848-1472 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 77-79 and ending at the stop codon at nucleotide positions 1445-1447 (Figure 166). The predicted polypeptide precursor is 456 amino acids long (Figure 167). The full-length PRO1108 protein shown in Figure 167 has an estimated molecular weight of about 52,071 daltons and a pI of about 9.46. Analysis of the full-length PRO1108 sequence shown in Figure 167 (SEQ ID NO:248) evidences the presence of the following: type II transmembrane domains from about amino acid 22 to about amino acid 42, from about amino acid 156 to about amino acid 176, from about amino acid 180 to about amino acid 199 and from about amino acid 369 to about amino acid 388, potential N-glycosylation sites from about amino acid 247 to about amino acid 250, from about amino acid 327 to about amino acid 330, from about amino acid 328 to about amino acid 331 and from about amino acid 362 to about amino acid 365 and an amino acid block having homology to ER lumen protein retaining receptor protein from about amino acid 153 to about amino acid 190. Clone DNA58848-1472 has been deposited with ATCC on June 9, 1998 and is assigned ATCC deposit no. 209955.

Analysis of the amino acid sequence of the full-length PRO1108 polypeptide suggests that it possesses significant sequence similarity to the LPAAT protein, thereby indicating that PRO1108 may be a novel LPAAT homolog. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO1108 amino acid sequence and the following Dayhoff sequences, AF015811_1, CER07E3_2, YL35_CAEEL, S73863, CEF59F4_4, P_W06422, MMU41736_1, MTV008_39, P_R99248 and Y67_BPT7.

EXAMPLE 73: Isolation of cDNA Clones Encoding Human PRO1137

The extracellular domain (ECD) sequences (including the secretion signal, if any) of from about 950 known secreted proteins from the Swiss-Prot public protein database were used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Using this procedure, Incyte EST No. 3459449, also referred to herein as "DNA7108", was identified as an EST having a BLAST score of 70 or greater that did not encode a known protein.

A consensus DNA sequence was assembled relative to the DNA7108 sequence and other ESTs using repeated cycles of BLAST and the program "phrap" (Phil Green, Univ. of Washington, Seattle, WA). The consensus sequence obtained therefrom is referred to herein as DNA53952.

In light of an observed sequence homology between the DNA53952 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3663102, the Incyte EST clone 3663102 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 168.

The entire nucleotide sequence of DNA58849-1494 is shown in Figure 168 (SEQ ID NO:249). Clone DNA58849-1494 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 77-79 and ending at the stop codon at nucleotide positions 797-799 (Figure 168). The predicted polypeptide precursor is 240 amino acids long (Figure 169). The full-length PRO1137 protein shown in Figure 169 has an estimated molecular weight of about 26,064 daltons and a pI of about 8.65. Analysis of the full-length PRO1137 sequence shown in Figure 169 (SEQ ID NO:250) evidences the presence of a signal peptide at about amino acids 1 to 14 and a potential N-glycosylation site at about amino acids 101-105.

Analysis of the amino acid sequence of the full-length PRO1137 polypeptide suggests that it possesses significant sequence similarity to ribosyltransferase thereby indicating that PRO1137 may be a novel member of the ribosyltransferase family of proteins. Analysis of the amino acid sequence of the full-length PRO1137 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1137 amino acid sequence and the following Dayhoff sequences: MMART5_1, NARG_MOUSE, GEN11909, GEN13794, GEN14406, MMRNART62_1, and P_R41876.

EXAMPLE 74: Isolation of cDNA clones Encoding Human PRO1138

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST sequence, Incyte cluster sequence no. 165212. This cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated as DNA54224. The assembly included a proprietary Genentech EST designated herein as DNA49140 (Figure 172; SEQ ID NO:254).

In light of an observed sequence homology between the DNA54224 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3836613, the Incyte EST clone 3836613 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 170 and is the full-length DNA sequence for PRO1138. Clone DNA58850-1495 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209956.

The entire nucleotide sequence of DNA58850-1495 is shown in Figure 170 (SEQ ID NO:252). Clone DNA58850-1495 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 38-40 and ending at the stop codon at nucleotide positions 1043-1045 (Figure 170). The predicted polypeptide precursor is 335 amino acids long (Figure 171). The full-length PRO1138 protein shown in Figure 171 has an estimated molecular weight of about 37,421 Daltons and a pI of about 6.36. Analysis of the full-length PRO1138 sequence shown in Figure 171 (SEQ ID NO:253) evidences the presence of the following features: a signal peptide at about amino acid 1 to about amino acid 22; a transmembrane domain at about amino

acids 224 to about 250; a leucine zipper pattern at about amino acids 229 to about 250; and potential N-glycosylation sites at about amino acids 98-101, 142-145, 148-151, 172-175, 176-179, 204-207, and 291-295.

Analysis of the amino acid sequence of the full-length PRO1138 polypeptide suggests that it possesses significant sequence similarity to the CD84, thereby indicating that PRO1138 may be a novel member of the Ig superfamily of polypeptides. More particularly, analysis of the amino acid sequence of the full-length PRO1138 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1138 amino acid sequence and the following Dayhoff sequences: HSU82988_1, HUMLY9_1, P_R97631, P_R97628, P_R97629, P_R97630, CD48_RAT, CD2_HUMAN, P_P93996, and HUMBGP_1.

Clone DNA58850-1495 was deposited with ATCC on June 9, 1998, and is assigned ATCC deposit no. 209956.

EXAMPLE 75: Isolation of cDNA clones Encoding Human PRO1054

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 66212. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55722.

In light of an observed sequence homology between the DNA55722 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 319751, the Incyte EST clone 319751 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 173 and is herein designated as DNA58853-1423.

Clone DNA58853-1423 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 46-48 and ending at the stop codon at nucleotide positions 586-588 (Figure 173). The predicted polypeptide precursor is 180 amino acids long (Figure 174). The full-length PRO1054 protein shown in Figure 174 has an estimated molecular weight of about 20,638 daltons and a pI of about 5.0. Analysis of the full-length PRO1054 sequence shown in Figure 174 (SEQ ID NO:256) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 18, a leucine zipper pattern from about amino acid 155 to about amino acid 176 and amino acid sequence blocks having homology to lipocalin proteins from about amino acid 27 to about amino acid 38 and from about amino acid 110 to about amino acid 120. Clone DNA58853-1423 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203016.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 174 (SEQ ID NO:256), evidenced significant homology between the PRO1054 amino acid sequence and the following Dayhoff sequences: MUP1_MOUSE, MUP6_MOUSE, MUP2_MOUSE, MUP8_MOUSE, MUP5_MOUSE, MUP4_MOUSE, S10124.

MUPM_MOUSE, MUP_RAT and ECU70823_1.

EXAMPLE 76: Isolation of cDNA clones Encoding Human PRO994

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 157555. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55728.

In light of an observed sequence homology between the DNA55728 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2860366, the Incyte EST clone 2860366 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.

The sequence of this cDNA insert is shown in Figure 175 and is herein designated as DNA58855-1422.

Clone DNA58855-1422 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 31-33 and ending at the stop codon at nucleotide positions 718-720 (Figure 175). The predicted polypeptide precursor is 229 amino acids long (Figure 176). The full-length PRO994 protein shown in Figure 176 has an estimated molecular weight of about 25,109 daltons and a pI of about 6.83. Analysis of the full-length PRO994 sequence shown in Figure 176 (SEQ ID NO:258) evidences the presence of the following: transmembrane domains from about amino acid 10 to about amino acid 31, from about amino acid 50 to about amino acid 72, from about amino acid 87 to about amino acid 110 and from about amino acid 191 to about amino acid 213, potential N-glycosylation sites from about amino acid 80 to about amino acid 83, from about amino acid 132 to about amino acid 135, from about amino acid 148 to about amino acid 151 and from about amino acid 163 to about amino acid 166 and an amino acid block having homology to TNFR/NGFR cysteine-rich region proteins from about amino acid 4 to about amino acid 11. Clone DNA58855-1422 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203018.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 176 (SEQ ID NO:258), evidenced significant homology between the PRO994 amino acid sequence and the following Dayhoff sequences: AF027204_1, TAL6_HUMAN, ILT4_HUMAN, JC6205, MMU57570_1, S40363, ETU56093_1, S42858, P_R66849 and P_R74751.

EXAMPLE 77: Isolation of cDNA clones Encoding Human PRO812

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 170079. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank)

and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated as DNA55721.

In light of an observed sequence homology between the DNA55721 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 388964, the Incyte EST clone 388964 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 177 and is herein designated as DNA59205-1421.

Clone DNA59205-1421 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 55-57 and ending at the stop codon at nucleotide positions 304-306 (Figure 177). The predicted polypeptide precursor is 83 amino acids long (Figure 178). The full-length PRO812 protein shown in Figure 178 has an estimated molecular weight of about 9,201 daltons and a pI of about 9.3. Analysis of the full-length PRO812 sequence shown in Figure 178 (SEQ ID NO:260) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 15, a cAMP- and cGMP-dependent protein kinase phosphorylation site from about amino acid 73 to about amino acid 76 and protein kinase C phosphorylation sites from about amino acid 70 to about amino acid 72 and from about amino acid 76 to about amino acid 78. Clone DNA59205-1421 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203009.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 178 (SEQ ID NO:260), evidenced significant homology between the PRO812 amino acid sequence and the following Dayhoff sequences: P_W35802, P_W35803, PSC1_RAT, S68231, GEN13917, PSC2_RAT, CC10_HUMAN, UTER_RABBIT, AF008595_1 and A56413.

EXAMPLE 78: Isolation of cDNA clones Encoding Human PRO1069

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST sequence designated herein as 100727. This sequence was then compared to a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, Univ. of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56001.

In light of an observed sequence homology between the DNA56001 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3533881, the Incyte EST clone 3533881 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 179 and is the full-length DNA sequence for PRO1069.

Clone DNA59211-1450 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209960.

The entire nucleotide sequence of DNA59211-1450 is shown in Figure 179 (SEQ ID NO:261). Clone DNA59211-1450 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 197-199 and ending at the stop codon at nucleotide positions 464-466. The predicted polypeptide precursor is 89 amino acids long (Figure 180). The full-length PRO1069 protein shown in Figure 180 has an estimated molecular weight of about 9,433 daltons and a pI of about 8.21. Analysis of the full-length PRO1069 sequence shown in Figure 180 (SEQ ID NO:262) evidences the presence of the following features: a signal peptide sequence at amino acid 1 to about 16; a transmembrane domain at about amino acids 36 to about 59; potential N-myristoylation sites at about amino acids 41-46, 45-50, and 84-89; and homology with extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 at about amino acids 54 to about 66.

Analysis of the amino acid sequence of the full-length PRO1069 polypeptide suggests that it possesses significant sequence similarity to CHIF, thereby indicating that PRO1069 may be a member of the CHIF family of polypeptides. More particularly, analysis of the amino acid sequence of the full-length PRO1069 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1069 amino acid sequence and the following Dayhoff sequences: CHIF_RAT, A55571, PLM_HUMAN, A40533, ATNG_BOVIN, RIC_MOUSE, PETD_SYNY3, VTBI_XENLA, A05009, and S75086.

Clone DNA59211-1450 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209960.

EXAMPLE 79: Isolation of cDNA Clones Encoding Human PRO1129

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence designated herein as 98833. The Incyte EST cluster sequence no. 98833 sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56038.

In light of an observed sequence homology between the DNA56038 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1335241, the Incyte EST clone 1335241 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 181 and is herein designated DNA59213-1487.

The full length clone shown in Figure 181 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 42-44 and ending at the stop codon found at nucleotide positions 1614-1616 (Figure 181; SEQ ID NO:263). The predicted polypeptide precursor is 524 amino acids long, has a calculated molecular weight of approximately 60,310 daltons and an estimated pI of approximately 7.46.

Analysis of the full-length PRO1129 sequence shown in Figure 182 (SEQ ID NO:264) evidences the presence of the following: type II transmembrane domains from about amino acid 13 to about amino acid 32 and from about amino acid 77 to about amino acid 102, a cytochrome P-450 cysteine heme-iron ligand signature sequence from about amino acid 461 to about amino acid 470 and potential N-glycosylation sites from about amino acid 112 to about amino acid 115 and from about amino acid 168 to about amino acid 171. Clone DNA59213-1487 has been deposited with the ATCC on June 9, 1998 and is assigned ATCC deposit no. 209959.

Analysis of the amino acid sequence of the full-length PRO1129 polypeptide suggests that it possesses sequence similarity to the cytochrome P-450 family of proteins. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some degree of homology between the PRO1129 amino acid sequence and the following Dayhoff sequences, AC004523_1, S45702, AF054821_1 and I53015.

EXAMPLE 80: Isolation of cDNA clones Encoding Human PRO1068

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the LIFESEQ® database, designated Incyte cluster no. 141736. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. One or more of the ESTs was derived from a human mast cell line from a patient with mast cell leukemia. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56094.

In light of an observed sequence homology between the DNA56094 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 004974, the Incyte EST clone 004974 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 183 and is herein designated as DNA59214-1449 (SEQ ID NO:265).

The full length clone shown in Figure 183 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 42-44 and ending at the stop codon found at nucleotide positions 414-416 (Figure 183; SEQ ID NO:265). The predicted polypeptide precursor (Figure 184, SEQ ID NO:266) is 124 amino acids long. PRO1068 has a calculated molecular weight of approximately 14,284 daltons and an estimated pI of approximately 8.14. The PRO1068 polypeptide has the following additional features, as indicated in Figure 184: a signal peptide sequence at about amino acids 1-20, a urotensin II signature sequence at about amino acids 118-123, a cell attachment sequence at about amino acids 64-66, and a potential cAMP- and cGMP-dependent protein kinase phosphorylation site at about amino acids 112-115.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 184 (SEQ ID NO:266), revealed homology between the PRO1068 amino acid sequence and the following Dayhoff sequences: HALBOP_1, MTVO43_36,

150498, and P_R78445

Clone DNA59214-1449 was deposited with the ATCC on July 1, 1998 and is assigned ATCC deposit no.203046.

EXAMPLE 81: Isolation of cDNA clones Encoding Human PRO1066

5 Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence designated herein as 79066. The Incyte EST cluster sequence no. 79066 sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or
10 BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56121.

In light of an observed sequence homology between the DNA56121 consensus sequence and an EST
15 sequence encompassed within the Incyte EST clone no. 1515315, the Incyte EST clone 1515315 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 185 and is herein designated DNA59215-1425.

The full length clone shown in Figure 185 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 176-178 and ending at the stop codon found at nucleotide
20 positions 527-529 (Figure 185; SEQ ID NO:267). The predicted polypeptide precursor is 117 amino acids long, has a calculated molecular weight of approximately 12,911 daltons and an estimated pI of approximately 5.46. Analysis of the full-length PRO1066 sequence shown in Figure 186 (SEQ ID NO:268) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 23, a cAMP- and cGMP-dependent protein kinase phosphorylation site from about amino acid 38 to about amino acid 41 and potential
25 N-myristoylation sites from about amino acid 5 to about amino acid 10, from about amino acid 63 to about amino acid 68 and from about amino acid 83 to about amino acid 88. Clone UNQ524 (DNA59215-1425) has been deposited with the ATCC on June 9, 1998 and is assigned ATCC deposit no. 209961.

Analysis of the amino acid sequence of the full-length PRO1066 polypeptide suggests that it does not possess significant sequence similarity to any known human protein. However, an analysis of the Dayhoff
30 database (version 35.45 SwissProt 35) evidenced some degree of homology between the PRO1066 amino acid sequence and the following Dayhoff sequences, MOTI_HUMAN, AF025667_1, MTCY19H9_8 and RABIGKCH_1.

EXAMPLE 82: Isolation of cDNA Clones Encoding Human PRO1184

35 Use of the signal sequence algorithm described in Example 3 on ESTs from an Incyte database allowed identification a candidate sequence designated herein as DNA56375. This sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and

a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56375.

In light of an observed sequence homology between the DNA56375 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1428374, the Incyte EST clone 1428374 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 187.

The full length clone shown in Figure 187 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 106-108 and ending at the stop codon found at nucleotide positions 532-534 (Figure 187; SEQ ID NO:269). The predicted polypeptide precursor is 142 amino acids long, has a calculated molecular weight of approximately 15,690 daltons and an estimated pI of approximately 9.64. Analysis of the full-length PRO1184 sequence shown in Figure 188 (SEQ ID NO:270) evidences the presence of a signal peptide at about amino acids 1-38. Clone DNA59220-1514 has been deposited with the ATCC on June 9, 1998. It is understood that the deposited clone has the actual sequences and that representations are presented herein.

Analysis of the amino acid sequence of the full-length PRO1184 polypeptide suggests that it possesses some sequence identity with a protein called TIM from *Drosophila virilis*, designated "DVTIMS02_1" in the Dayhoff data base, (version 35.45 SwissProt 35). Other Dayhoff database (version 35.45 SwissProt 35) sequences having some degree of sequence identity with PRO1184 include: WIS1_SCHPO, F002186_1, ATAC00239124 and MSAIPRP_1.

EXAMPLE 83: Isolation of cDNA clones Encoding Human PRO1360

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST sequence from an Incyte database, designated DNA10572. This EST sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank, Merck/Wash. U.) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57314.

In light of an observed sequence homology between the DNA57314 consensus sequence and an EST sequence encompassed within the Merck EST clone no. AA406443, the Merck EST clone AA406443 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 189 and is herein designated as DNA59488-1603.

The full length clone shown in Figure 189 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 54-56 and ending at the stop codon found at nucleotide positions 909-911 (Figure 189; SEQ ID NO:271). The predicted polypeptide precursor (Figure 190, SEQ ID NO:272) is 285 amino acids long. PRO1360 has a calculated molecular weight of approximately 31,433 daltons and an estimated pI of approximately 7.32. Clone DNA59488-1603 was deposited with the ATCC on August 25, 1998 and is assigned ATCC deposit no. 203157.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 190 (SEQ ID NO:272), revealed sequence identity between the PRO1360 amino acid sequence and the following Dayhoff sequences: UN51_CAEEL, YD4B_SCHPO, AF000634_1, GFO_ZYMMO, YE1J_SCHPO, D86566_1, ZMGFO_1, S76976, PPSA_SYNY3, and CEF28B1_4.

EXAMPLE 84: Isolation of cDNA clones Encoding Human PRO1029

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 18763. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57854.

In light of an observed sequence homology between the DNA57854 consensus sequence and an EST sequence encompassed within the Merck EST clone no. T98880, the Merck EST clone T98880 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 191 and is herein designated as DNA59493-1420.

Clone DNA59493-1420 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 39-41 and ending at the stop codon at nucleotide positions 297-299 (Figure 191). The predicted polypeptide precursor is 86 amino acids long (Figure 192). The full-length PRO1029 protein shown in Figure 192 has an estimated molecular weight of about 9,548 daltons and a pI of about 8.52. Analysis of the full-length PRO1029 sequence shown in Figure 192 (SEQ ID NO:274) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 19, an amino acid block having homology to bacterial rhodopsins retinal binding site protein from about amino acid 50 to about amino acid 61, a prenyl group binding site from about amino acid 83 to about amino acid 86 and a potential N-glycosylation site from about amino acid 45 to about amino acid 48. Clone DNA59493-1420 has been deposited with ATCC on July 1, 1998 and is assigned ATCC deposit no. 203050.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 192 (SEQ ID NO:274), evidenced significant

homology between the PRO1029 amino acid sequence and the following Dayhoff sequences: S66088, AF031815_1, MM4A6L_1, PSEIS52a-1, S17699 and P_R63635.

EXAMPLE 85: Isolation of cDNA clones Encoding Human PRO1139

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 4461. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57312.

The DNA57312 consensus sequence included a 172 nucleotides long public EST (T62095, Merck/University of Washington public database). This EST clone, identified herein as a putative protein coding sequence, was purchased from Merck, and sequenced to provide the coding sequence of PRO1139 (Figure 193). As noted before, the deduced amino acid sequence of DNA59497-1496 shows a significant sequence identity with the deduced amino acid sequence of HSOBRGRP_1. The full-length protein (Figure 194) contains a putative signal peptide between amino acid residues 1 and about 28, and three putative transmembrane domains (approximate amino acid residues 33-52, 71-89, 98-120).

EXAMPLE 86: Isolation of cDNA clones Encoding Human PRO1309

An expressed sequence tag (EST) DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST was identified which showed homology to SLIT.

RNA for construction of cDNA libraries was isolated from human fetal brain tissue. The cDNA libraries used to isolate the cDNA clones encoding human PRO1309 were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI.

The cDNA libraries (prepared as described above), were screened by hybridization with a synthetic oligonucleotide probe derived from the above described Incyte EST sequence:

5'-TCCGTGCAGGGGGACGCCTTTCAGAACTGCGCCGAGTTAAGGAAC-3' (SEQ ID NO:279).

A cDNA clone was isolated and sequenced in entirety. The entire nucleotide sequence of DNA59588-1571 is shown in Figure 195 (SEQ ID NO:277). Clone DNA59588-1571 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 720-722 and a stop codon at nucleotide positions 2286-2288 (Figure 195; SEQ ID NO:277). The predicted polypeptide precursor is 522 amino acids

long. The signal peptide is approximately at 1-34 and the transmembrane domain is at approximately 428-450 of SEQ ID NO:278. Clone DNA59588-1571 has been deposited with ATCC and is assigned ATCC deposit no. 203106. The full-length PRO1309 protein shown in Figure 196 has an estimated molecular weight of about 58,614 daltons and a pI of about 7.42.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 196 (SEQ ID NO:278), revealed sequence identity between the PRO1309 amino acid sequence and the following Dayhoff sequences: AB007876_1, GPV_MOUSE, ALS_RAT, P_R85889, LUM_CHICK, AB014462_1, PGS1_CANFA, CEM88_7, A58532 and GEN11209.

EXAMPLE 87: Isolation of cDNA Clones Encoding Human PRO1028

Use of the signal sequence algorithm described in Example 3 above allowed identification of a certain EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA59603.

In light of an observed sequence homology between the DNA59603 sequence and an EST sequence contained within Incyte EST clone no. 1497725, the Incyte EST clone no. 1497725 was purchased and the cDNA insert was obtained and sequenced. It was found that the insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 197 and is herein designated as DNA59603-1419.

The entire nucleotide sequence of DNA59603-1419 is shown in Figure 197 (SEQ ID NO:280). Clone DNA59603-1419 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 21-23 and ending at the stop codon at nucleotide positions 612-614 (Figure 197). The predicted polypeptide precursor is 197 amino acids long (Figure 198). The full-length PRO1028 protein shown in Figure 198 has an estimated molecular weight of about 20,832 daltons and a pI of about 8.74. Clone DNA59603-1419 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analyzing the amino acid sequence of SEQ ID NO:281, the putative signal peptide is at about amino acids 1-19 of SEQ ID NO:281. An N-glycosylation site is at about amino acids 35-38 of SEQ ID NO:281. A C-type lectin domain is at about amino acids 108-117 of SEQ ID NO:281, indicating that PRO513 may be related to or be a lectin. The corresponding nucleotides of these amino acid sequences or others can be routinely determined given the sequences provided herein.

EXAMPLE 88: Isolation of cDNA Clones Encoding Human PRO1027

Use of the signal sequence algorithm described in Example 3 above allowed identification of a certain EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56399.

In light of an observed sequence homology between the DNA56399 sequence and an EST sequence contained within Incyte EST clone no. 937605, the Incyte EST clone no. 937605 was purchased and the cDNA insert was obtained and sequenced. It was found that the insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 199 and is herein designated as DNA59605-1418.

The entire nucleotide sequence of DNA59605-1418 is shown in Figure 199 (SEQ ID NO:282). Clone DNA59605-1418 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 31-33 and ending at the stop codon at nucleotide positions 262-264 (Figure 199). The predicted polypeptide precursor is 77 amino acids long (Figure 200). The full-length PRO1027 protein shown in Figure 200 has an estimated molecular weight of about 8,772 daltons and a pI of about 9.62. Clone DNA59605-1418 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analyzing the amino acid sequence of SEQ ID NO:283, the putative signal peptide is at about amino acids 1-33 of SEQ ID NO:283. The type II fibronectin collagen-binding domain begins at about amino acid 30 of SEQ ID NO:283. The corresponding nucleotides for these amino acid sequences and others can be routinely determined given the sequences provided herein. PRO1027 may be involved in tissue formation or repair.

The following Dayhoff designations appear to have some sequence identity with PRO1027: SFT2_YEAST; ATM3E9_2; A69826; YM16_MARPO; E64896; U60193_2; MTLRAJ205_1; MCU60315_70; SPAS_SHIFL; and S54213.

EXAMPLE 89: Isolation of cDNA Clones Encoding Human PRO1107

Use of the signal sequence algorithm described in Example 3 above allowed identification of a certain EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence

obtained therefrom is herein designated DNA56402.

In light of an observed sequence homology between the DNA56402 sequence and an EST sequence contained within Incyte EST clone no. 3203694, the Incyte EST clone no. 3203694 was purchased and the cDNA insert was obtained and sequenced. It was found that the insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 201 and is herein designated as DNA59606-1471.

- 5 The entire nucleotide sequence of DNA59606-1471 is shown in Figure 201 (SEQ ID NO:284). Clone DNA59606-1471 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 244-246 and ending at the stop codon at nucleotide positions 1675-1677 of SEQ ID NO:284 (Figure 201). The predicted polypeptide precursor is 477 amino acids long (Figure 202). The full-length PRO1107 protein shown in Figure 202 has an estimated molecular weight of about 54,668 daltons and a pI of about 6.33.
- 10 Clone DNA59606-1471 has been deposited with ATCC on June 9, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

- Analysis of the amino acid sequence of the full-length PRO1107 polypeptide suggests that it possesses significant sequence similarity to phosphodiesterase I/nucleotide pyrophosphatase, human insulin receptor
- 15 tyrosine kinase inhibitor, alkaline phosphodiesterase and autotaxin, thereby indicating that PRO1107 may have at least one or all of the activities of these proteins, and that PRO1107 is a novel phosphodiesterase. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced sequence identity between the PRO1107 amino acid sequence and at least the following Dayhoff sequences: AF005632_1, P_R79148, RNU78787_1, AF060218_4, A57080 and HUMATXT_1.

20

EXAMPLE 90: Isolation of cDNA clones Encoding Human PRO1140

- Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST sequence, Incyte cluster sequence No. 135917. This sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary
- 25 EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, Univ. of Washington, Seattle, Washington). The consensus sequence obtained
- 30 therefrom is herein designated DNA56416.

- In light of an observed sequence homology between DNA56416 and an EST sequence contained within Incyte EST clone no. 3345705, Incyte EST clone no. 3345705 was obtained and its insert sequenced. It was found that the insert encoded a full-length protein. The sequence, designated herein as DNA59607-1497, which is shown in Figure 203, is the full-length DNA sequence for PRO1140. Clone DNA59607-1497 was deposited
- 35 with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209946.

The entire nucleotide sequence of DNA59607-1497 is shown in Figure 203 (SEQ ID NO:286). Clone DNA59607-1497 contains a single open reading frame with an apparent translational initiation site at nucleotide

positions 210-212 and ending at the stop codon at nucleotide positions 975-977 (Figure 203). The predicted polypeptide precursor is 255 amino acids long (Figure 204). The full-length PRO1140 protein shown in Figure 204 has an estimated molecular weight of about 29,405 daltons and a pI of about 7.64. Analysis of the full-length PRO1140 sequence shown in Figure 204 (SEQ ID NO:287) evidences the presence of three transmembrane domains at about amino acids 101 to 118, 141 to 161 and 172 to 191.

5 Analysis of the amino acid sequence of the full-length PRO1140 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1140 amino acid sequence and the following Dayhoff sequences: AF023602_1, AF000368_1, CIN3_RAT, AF003373_1, GEN13279, and AF003372_1.

Clone DNA59607-1497 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209946.

10

EXAMPLE 91: Isolation of cDNA clones Encoding Human PRO1106

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST sequence. This sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using
15 the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, Univ. of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated
20 DNA56423.

In light of an observed sequence homology between DNA56423 and an EST sequence contained within Incyte EST clone no. 1711247, Incyte EST clone no. 1711247 was obtained and its insert sequenced. It was found that the insert encoded a full-length protein. The sequence, designated herein as DNA59609-1470, which is shown in Figure 205, is the full-length DNA sequence for PRO1106. Clone DNA59609-1470 was deposited
25 with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209963.

The entire nucleotide sequence of DNA59609-1470 is shown in Figure 205 (SEQ ID NO:288). Clone DNA59609-1470 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 61-63 and ending at the stop codon at nucleotide positions 1468-1470 of SEQ ID NO:288 (Figure 205). The predicted polypeptide precursor is 469 amino acids long (Figure 206). The full-length PRO1106 protein
30 shown in Figure 206 has an estimated molecular weight of about 52,689 daltons and a pI of about 8.68. It is understood that the skilled artisan can construct the polypeptide or nucleic acid encoding therefor to exclude any one or more of all of these domains. For example, the transmembrane domain region(s) and/or either of the amino terminal or carboxyl end can be excluded. Clone DNA59609-1470 has been deposited with ATCC on June 9, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the
35 sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1106 polypeptide suggests that it possesses significant sequence similarity to the peroxisomal ca-dependent solute carrier, thereby indicating that PRO1106

may be a novel transporter. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced sequence identity between the PRO1106 amino acid sequence and at least the following Dayhoff sequences, AF004161_1, IG002N01_25, GDC_BOVIN and BT1_MAIZE.

EXAMPLE 92: Isolation of cDNA clones Encoding Human PRO1291

5 Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 120480. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul
10 et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56425.

In light of an observed sequence homology between the DNA56425 sequence and an EST sequence
15 encompassed within the Incyte EST clone no. 2798803, the Incyte EST clone 2798803 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 207 and is herein designated as DNA59610-1556.

Clone DNA59610-1556 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 61-63 and ending at the stop codon at nucleotide positions 907-909 (Figure 207). The
20 predicted polypeptide precursor is 282 amino acids long (Figure 208). The full-length PRO1291 protein shown in Figure 208 has an estimated molecular weight of about 30,878 daltons and a pI of about 5.27. Analysis of the full-length PRO1291 sequence shown in Figure 208 (SEQ ID NO:291) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 28, a transmembrane domain from about amino acid 258 to about amino acid 281 and potential N-glycosylation sites from about amino acid 112 to about
25 amino acid 115, from about amino acid 160 to about amino acid 163, from about amino acid 190 to about amino acid 193, from about amino acid 196 to about amino acid 199, from about amino acid 205 to about amino acid 208, from about amino acid 216 to about amino acid 219 and from about amino acid 220 to about amino acid 223.. Clone DNA59610-1556 has been deposited with ATCC on June 16, 1998 and is assigned ATCC deposit no. 209990.

30 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 208 (SEQ ID NO:291), evidenced significant homology between the PRO1291 amino acid sequence and the following Dayhoff sequences: HSU90552_1, HSU90144_1, AF033107_1, HSB73_1, HSU90142_1, GGCD80_1, P_W34452, MOG_MOUSE, B39371 and P_R71360.

35

EXAMPLE 93: Isolation of cDNA clones Encoding Human PRO1105

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56430.

In light of an observed sequence homology between the DNA56430 sequence and an EST sequence encompassed within the Incyte EST clone no. 1853047, the Incyte EST clone 1853047 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 209 and is herein designated as DNA59612-1466.

The entire nucleotide sequence of DNA59612-1466 is shown in Figure 209 (SEQ ID NO:292). Clone DNA59612-1466 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 28-30 and ending at the stop codon at nucleotide positions 568-570 of SEQ ID NO:292 (Figure 209). The predicted polypeptide precursor is 180 amino acids long (Figure 210). The full-length PRO1105 protein shown in Figure 210 has an estimated molecular weight of about 20,040 daltons and a pI of about 8.35. Clone DNA59612-1466 has been deposited with the ATCC on June 9, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analyzing Figure 210, a signal peptide is at about amino acids 1-19 of SEQ ID NO:293 and transmembrane domains are shown at about amino acids 80-99 and 145-162 of SEQ ID NO:293. It is understood that the skilled artisan could form a polypeptide with all of or any combination or individual selection of these regions. It is also understood that the corresponding nucleic acids can be routinely identified and prepared based on the information provided herein.

EXAMPLE 94: Isolation of cDNA clones Encoding Human PRO511

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56434.

In light of an observed sequence homology between the DNA56434 sequence and an EST sequence encompassed within the Incyte EST clone no. 1227491, the Incyte EST clone 1227491 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 211 and is herein designated as DNA59613-1417.

The entire nucleotide sequence of DNA59613-1417 is shown in Figure 211 (SEQ ID NO:294). Clone DNA59613-1417 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 233-235 and ending at the stop codon at nucleotide positions 944-946 (Figure 211). The predicted polypeptide precursor is 237 amino acids long (Figure 212). The full-length PRO511 protein shown in Figure 212 has an estimated molecular weight of about 25,284 daltons and a pI of about 5.74. Clone DNA59613-1417 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analyzing the amino acid sequence of SEQ ID NO:295, the putative signal peptide is at about amino acids 1-25 of SEQ ID NO:295. The N-glycosylation sites are at about amino acids 45-48, 73-76, 107-110, 118-121, 132-135, 172-175, 175-178 and 185-188 of SEQ ID NO:295. An arthropod defensins conserved region is at about amino acids 176-182 of SEQ ID NO:295. A kringle domain begins at about amino acid 128 of SEQ ID NO:295 and a ly-6/u-PAR domain begins at about amino acid 6 of SEQ ID NO:295. The corresponding nucleotides of these amino acid sequences and others can be routinely determined given the sequences provided herein.

The designations appearing in a Dayhoff database with which PRO511 has some sequence identity are as follows: SSC20F10_1; SF041083; P_W26579; S44208; JC2394; PSTA_DICDI; A27020; S59310; RAG1_RABIT; and MUSBALBC1_1.

EXAMPLE 95: Isolation of cDNA clones Encoding Human PRO1104

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56446.

In light of an observed sequence homology between the DNA56446 sequence and an EST sequence encompassed within the Incyte EST clone no. 2837496, the Incyte EST clone 2837496 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 213 and is herein designated as DNA59616-1465.

The entire nucleotide sequence of DNA59616-1465 is shown in Figure 213 (SEQ ID NO:296). Clone DNA59616-1465 contains a single open reading frame with an apparent translational initiation site at nucleotide

positions 109-111 and ending at the stop codon at nucleotide positions 1132-1134 of SEQ ID NO:296 (Figure 213). The predicted polypeptide precursor is 341 amino acids long (Figure 214). The full-length PRO1104 protein shown in Figure 214 has an estimated molecular weight of about 36,769 daltons and a pI of about 9.03. Clone DNA59616-1465 has been deposited with ATCC on June 16, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analyzing Figure 214, a signal peptide is at about amino acids 1-22 of SEQ ID NO:297. N-myristoylation sites are at about amino acids 41-46, 110-115, 133-138, 167-172 and 179-184 of SEQ ID NO:297.

10 EXAMPLE 96: Isolation of cDNA clones Encoding Human PRO1100

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

In light of an observed sequence homology between the obtained consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2305379, the Incyte EST clone 2305379 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 215 and is herein designated as DNA59619-1464.

The entire nucleotide sequence of DNA59619-1464 is shown in Figure 215 (SEQ ID NO:298). Clone DNA59619-1464 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 33-35 and ending at the stop codon at nucleotide positions 993-995 of SEQ ID NO:298 (Figure 215). The predicted polypeptide precursor is 320 amino acids long (Figure 216). The full-length PRO1100 protein shown in Figure 216 has an estimated molecular weight of about 36,475 daltons and a pI of about 7.29. Clone DNA59619-1464 has been deposited with ATCC on July 1, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Upon analyzing SEQ ID NO:299, the approximate locations of the signal peptide, the transmembrane domains, an N-glycosylation site, an N-myristoylation site, a CUB domain and an amiloride-sensitive sodium channel domain are present. It is believed that PRO1100 may function as a channel. The corresponding nucleic acids for these amino acids and others can be routinely determined given SEQ ID NO:299..

EXAMPLE 97: Isolation of cDNA clones Encoding Human PRO836

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained is herein designated DNA56453.

In light of an observed sequence homology between the DNA56453 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2610075, the Incyte EST clone 2610075 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 217 and is herein designated as DNA59620-1463.

The entire nucleotide sequence of DNA59620-1463 is shown in Figure 217 (SEQ ID NO:300). Clone DNA59620-1463 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 65-67 and ending at the stop codon at nucleotide positions 1448-1450 of SEQ ID NO:300 (Figure 217). The predicted polypeptide precursor is 461 amino acids long (Figure 218). The full-length PRO836 protein shown in Figure 218 has an estimated molecular weight of about 52,085 daltons and a pI of about 5.36. Analysis of the full-length PRO836 sequence shown in Figure 218 (SEQ ID NO:301) evidences the presence of the following: a signal peptide, N-glycosylation sites, N-myristoylation sites, a domain conserved in the YJL126w/YLR351c/yhcX family of proteins, and a region having sequence identity with SLS1. Clone DNA59620-1463 has been deposited with ATCC on June 16, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO836 polypeptide suggests that it possesses some sequence similarity to SLS1, thereby indicating that PRO836 may be involved in protein translocation of the ER. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some homology between the PRO836 amino acid sequence and at least the following Dayhoff sequences, S58132, SPBC3B9_1, S66714, CRU40057_1 and IMA_CAEEL.

EXAMPLE 98: Isolation of cDNA clones Encoding Human PRO1141

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 11873. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or

in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56518.

In light of an observed sequence homology between the DNA56518 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2679995, the Incyte EST clone 2679995 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 219 and is herein designated as DNA59625-1498.

Clone DNA59625-1498 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 204-206 and ending at the stop codon at nucleotide positions 945-947 (Figure 219). The predicted polypeptide precursor is 247 amino acids long (Figure 220). The full-length PRO1141 protein shown in Figure 220 has an estimated molecular weight of about 26,840 daltons and a pI of about 8.19. Analysis of the full-length PRO1141 sequence shown in Figure 220 (SEQ ID NO:303) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 19 and transmembrane domains from about amino acid 38 to about amino acid 57, from about amino acid 67 to about amino acid 83, from about amino acid 117 to about amino acid 139 and from about amino acid 153 to about amino acid 170. Clone DNA59625-1498 has been deposited with ATCC on June 16, 1998 and is assigned ATCC deposit no. 209992.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 220 (SEQ ID NO:303), evidenced significant homology between the PRO1141 amino acid sequence and the following Dayhoff sequences: CEVF36H2L_2, PCRB7PRJ_1, AB000506_1, LEU95008_1, MRU87980_15, YIGM_ECOLI, STU65700_1, GHU62778_1, CYST_SYNY3 and AF009567_1.

EXAMPLE 99: Isolation of cDNA clones Encoding Human PRO1132

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein as DNA35934. Based on the DNA35934 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1132.

PCR primers (forward and reverse) were synthesized:

forward PCR primer: 5'-TCCTGTGACCACCCCTCTAACACC-3' (SEQ ID NO:310) and

reverse PCR primer: 5'-CTGGAACATCTGCTGCCCAGATTC-3' (SEQ ID NO:311).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus sequence which had the following nucleotide sequence:

5'-GTCGGATGACAGCAGCAGCCGCATCATCAATGGATCCGACTGCGATATGC-3' (SEQ ID NO:312).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1132 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1132 and the derived protein sequence for PRO1132.

The entire nucleotide sequence of PRO1132 is shown in Figure 225 (SEQ ID NO:308). Clone DNA59767-1489 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 354-356 and a stop codon at nucleotide positions 1233-1235 (Figure 225; SEQ ID NO:308). The predicted polypeptide precursor is 293 amino acids long. The signal peptide is at about amino acids 1-22 and the histidine active site is at about amino acids 104-109 of SEQ ID NO:309. Clone DNA59767-1489 has been deposited with ATCC (having the actual sequence rather than representations based on sequencing techniques as presented herein) and is assigned ATCC deposit no. 203108. The full-length PRO1132 protein shown in Figure 226 has an estimated molecular weight of about 32,020 daltons and a pI of about 8.7.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 226 (SEQ ID NO:309), revealed sequence identity between the PRO1132 amino acid sequence and the following Dayhoff sequences: SSU76256_1, P_W10694, MMAE000663_6, AF013988_1, U66061_8, MMAE000665_2, MMAE00066415, MMAE00066414, MMAE000665_4 and MMAE00066412.

EXAMPLE 100: Isolation of cDNA clones Encoding Human NL7 (PRO1346)

A single EST sequence (#1398422) was found in the LIFESEQ[®] database as described in Example 1 above. This EST sequence was renamed as DNA45668. Based on the DNA45668 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for NL7.

PCR primers (forward and reverse) were synthesized:

forward PCR primer: 5'-CACACGTCCAACCTCAATGGGCAG-3' (SEQ ID NO:315)

reverse PCR primer: 5'-GACCAGCAGGGCCAAGGACAAGG-3' (SEQ ID NO:316)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA45668 sequence which had the following nucleotide sequence:

hybridization probe:

5'-GTTCTCTGAGATGAAGATCCGGCCGGTCCGGGAGTACCGCTTAG-3'

(SEQ ID NO:317)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the NL7 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from a human fetal kidney library (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for NL7 (designated herein as DNA59776-1600 [Figure 227, SEQ ID NO:313]) and the derived protein sequence for NL7 (PRO1346).

The entire coding sequence of NL7 (PRO1346) is shown in Figure 227 (SEQ ID NO:313). Clone DNA59776-1600 contains a single open reading frame with an apparent translational initiation site at nucleotide

positions 1-3 and an apparent stop codon at nucleotide positions 1384-1386. The predicted polypeptide precursor is 461 amino acids long. The protein contains an apparent type II transmembrane domain at amino acid positions from about 31 to about 50; fibrinogen beta and gamma chains C-terminal domain signature starting at about amino acid position 409, and a leucine zipper pattern starting at about amino acid positions 140, 147, 154 and 161, respectively. Clone DNA59776-1600 has been deposited with ATCC and is assigned ATCC deposit no. 203128. The full-length NL7 protein shown in Figure 228 has an estimated molecular weight of about 50,744 daltons and a pI of about 6.38.

Based on a WU-BLAST2 sequence alignment analysis (using the WU-BLAST2 computer program) of the full-length sequence, NL7 shows significant amino acid sequence identity to a human microfibril-associated glycoprotein (1 MFA4_HUMAN); to known TIE-2 ligands and ligand homologues, ficolin, serum lectin and TGF-1 binding protein.

EXAMPLE 101: Isolation of cDNA clones Encoding Human PRO1131

A cDNA sequence isolated in the amylase screen described in Example 2 above is herein designated DNA43546 (see Figure 231; SEQ ID NO:320). The DNA43546 sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA45627.

Based on the DNA45627 sequence, oligonucleotide probes were generated and used to screen a human library prepared as described in paragraph 1 of Example 2 above. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp.

PCR primers (forward and 2 reverse) were synthesized:

forward PCR primer 5'-ATGCAGGCCAAGTACAGCAGCAC-3' (SEQ ID NO:321);

reverse PCR primer 1 5'-CATGCTGACGACTTCCTGCAAGC-3' (SEQ ID NO:322); and

reverse PCR primer 1 5'-CCACACAGTCTCTGCTTCTTGGG-3' (SEQ ID NO:323)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA45627 sequence which had the following nucleotide sequence:

hybridization probe

5'-ATGCTGGATGATGATGGGGACACCACCATGAGCCTGCATT-3' (SEQ ID NO:324).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1131 gene using the probe oligonucleotide and one of the PCR primers.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 144-146, and a stop signal at nucleotide positions 984-986 (Figure 229; SEQ ID NO:318). The predicted polypeptide precursor is 280 amino acids long, has a calculated molecular weight of approximately 31,966 daltons and an estimated pI of approximately 6.26. The transmembrane domain sequence is at about 49-74 of SEQ ID NO:319 and the region having sequence identity with LDL receptors is about 50-265 of SEQ ID NO:319. PRO1131 contains potential N-linked glycosylation sites at amino acid positions 95-98 and 169-172 of SEQ ID NO:319. Clone DNA59777-1480 has been deposited with the ATCC and is assigned ATCC deposit no. 203111.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 230 (SEQ ID NO:319), evidenced some sequence identity between the PRO1131 amino acid sequence and the following Dayhoff sequences: AB010710_1, I49053, I49115, RNU56863_1, LY4A_MOUSE, I55686, MMU56404_1, I49361, AF030313_1 and MMU09739_1.

EXAMPLE 102: Isolation of cDNA clones Encoding Human PRO1281

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein as DNA35720. Based on the DNA35720 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1281.

PCR primers (forward and reverse) were synthesized:

forward PCR primers:

5'-TGGAAGGCTGCCGCAACGACAATC-3' (SEQ ID NO:327);
5'-CTGATGTGGCCGATGTTCTG-3' (SEQ ID NO:328); and
5'-ATGGCTCAGTGTGCAGACAG-3' (SEQ ID NO:329).

reverse PCR primers:

5'-GCATGCTGCTCCGTGAAGTAGTCC-3' (SEQ ID NO:330); and
5'-ATGCATGGGAAAGAAGGCCTGCCC-3' (SEQ ID NO:331).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA35720 sequence which had the following nucleotide sequence:

hybridization probe:

5'-TGCACTGGTGACCACGAGGGGGTGCCTATAGCCATCTGGAGCTGAG-3' (SEQ ID NO:332).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO1281 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated human fetal liver.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1281 (designated herein as DNA59820-1549 [Figure 232, SEQ ID NO:325]; and the derived protein sequence for PRO1281.

The entire coding sequence of PRO1281 is shown in Figure 232 (SEQ ID NO:325). Clone DNA59820-1549 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 228-230 and an apparent stop codon at nucleotide positions 2553-2555. The predicted polypeptide precursor is 775 amino acids long. The full-length PRO1281 protein shown in Figure 233 has an estimated molecular weight of about 85,481 daltons and a pI of about 6.92. Additional features include a signal peptide at about amino acids 1-15; and potential N-glycosylation sites at about amino acids 138-141 and 361-364.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 233 (SEQ ID NO:326), revealed some sequence identity between the PRO1281 amino acid sequence and the following Dayhoff sequences: S44860, CET24D1_1, CEC38H2_3, CAC2_HAECO, B3A2_HUMAN, S22373, CEF38A3_2, CEC34F6_2, CEC34F6_3, and CELT22B11_3.

Clone DNA59820-1549 has been deposited with ATCC and is assigned ATCC deposit no: 203129.

EXAMPLE 103: Isolation of cDNA clones Encoding Human PRO1064

A cDNA sequence isolated in the amylase screen described in Example 2 above was found, by the WU-BLAST2 sequence alignment computer program, to have no significant sequence identity to any known human protein. This cDNA sequence is herein designated DNA45288. The DNA45288 sequence was then compared to various EST databases including public EST databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify homologous EST sequences. The comparison was performed using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)]. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). This consensus sequence is herein designated DNA48609. Oligonucleotide primers based upon the DNA48609 sequence were then synthesized and employed to screen a human fetal kidney cDNA library which resulted in the identification of the DNA59827-1426 clone shown in Figure 234. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp.

The oligonucleotide probes employed were as follows:

forward PCR primer 5'-CTGAGACCCTGCAGCACCATCTG-3' (SEQ ID NO:336)

reverse PCR primer 5'-GGTGCTTCTTGAGCCCCACTTAGC-3' (SEQ ID NO:337)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA48609 sequence which had the following nucleotide sequence

hybridization probe

5'-AATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCGCTGT-3' (SEQ ID NO:338)

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 532-534 and a stop signal at nucleotide positions 991-993 (Figure 234, SEQ ID NO:333). The predicted polypeptide precursor is 153 amino acids long, has a calculated

molecular weight of approximately 17,317 daltons and an estimated pI of approximately 5.17. Analysis of the full-length PRO1064 sequence shown in Figure 235 (SEQ ID NO:334) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 24, a transmembrane domain from about amino acid 89 to about amino acid 110, an indole-3-glycerol phosphate synthase homology block from about amino acid 74 to about amino acid 105 and a Myb DNA binding domain protein repeat protein homology block from about amino acid 114 to about amino acid 137. Clone DNA59827-1426 has been deposited with ATCC on August 4, 1998 and is assigned ATCC deposit no. 203089.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 235 (SEQ ID NO:334), evidenced homology between the PRO1064 amino acid sequence and the following Dayhoff sequences: MMNP15PRO_1, BP187PLYH_1, CELF42G8_4, MMU58888_1, GEN14270, TUB8_SOLTU, RCN_MOUSE, HUMRBSY79_1, SESENODA_1 and A21467_1.

EXAMPLE 104: Isolation of cDNA clones Encoding Human PRO1379

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein DNA45232. Based on the DNA45232 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1379.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGGACACCGTACCCTGGTATCTGC-3' (SEQ ID NO:341)

reverse PCR primer 5'-CCAACTCTGAGGAGAGCAAGTGGC-3' (SEQ ID NO:342)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA45232 sequence which had the following nucleotide sequence:

hybridization probe

5'-TGTATGTGCACACCCTCACCATCACCTCCAAGGGCAAGGAGAAC-3' (SEQ ID NO:343).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1379 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1379 which is designated herein as DNA59828-1608 and shown in Figure 237 (SEQ ID NO:339); and the derived protein sequence for PRO1379 (SEQ ID NO:340).

The entire coding sequence of PRO1379 is shown in Figure 237 (SEQ ID NO:339). Clone DNA59828-1608 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 10-12 and an apparent stop codon at nucleotide positions 1732-1734. The predicted polypeptide precursor is 574 amino acids long. The full-length PRO1379 protein shown in Figure 238 has an estimated molecular weight of about 65,355 daltons and a pI of about 8.73. Additional features include a signal peptide at about amino acids

1-17 and potential N-glycosylation sites at about amino acids 160-163, 287-290, and 323-326.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 238 (SEQ ID NO:340), revealed some homology between the PRO1379 amino acid sequence and the following Dayhoff sequences: YHY8_YEAST, AF040625_1, HP714394_1, and HIV18U45630_1.

5 Clone DNA59828-1608 has been deposited with ATCC and is assigned ATCC deposit no. 203158.

EXAMPLE 105: Isolation of cDNA Clones Encoding Human PRO844

10 An expressed sequence tag (EST) DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST was identified which showed sequence identity with aLP. Based on the information and discoveries provided herein, the clone for this EST, Incyte clone no. 2657496 from a cancerous lung library was further examined.

DNA sequencing of the insert for this clone gave a sequence (herein designated as DNA59838-1462; SEQ ID NO:344) which includes the full-length DNA sequence for PRO844 and the derived protein sequence for PRO844.

15 The entire nucleotide sequence of DNA59838-1462 is shown in Figure 239 (SEQ ID NO:344). Clone DNA59838-1462 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 5-7 and ending at the stop codon at nucleotide positions 338-340 of SEQ ID NO:344 (Figure 239). The predicted polypeptide precursor is 111 amino acids long (Figure 240). The full-length PRO844 protein shown in Figure 240 has an estimated molecular weight of about 12,050 daltons and a pI of about 5.45. Clone 20 UNQ544 DNA59838-1462 has been deposited with ATCC on June 16, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

25 Analysis of the amino acid sequence of the full-length PRO844 polypeptide suggests that it possesses significant sequence similarity to serine protease inhibitors, thereby indicating that PRO844 may be a novel proteinase inhibitor. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO844 amino acid sequence and at least the following Dayhoff sequences, ALK1_HUMAN, P_P82403, P_P82402, ELAF_HUMAN and P_P60950.

EXAMPLE 106: Isolation of cDNA Clones Encoding Human PRO848

30 Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in 35 Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence

obtained therefrom is herein designated DNA55999.

In light of an observed sequence homology between the DNA55999 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2768571, the Incyte EST clone 2768571 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 241 and is herein designated as DNA59839-1461.

5 The entire nucleotide sequence of DNA59839-1461 is shown in Figure 241 (SEQ ID NO:346). Clone DNA59839-1461 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 146-148 and ending at the stop codon at nucleotide positions 1946-1948 of SEQ ID NO:346 (Figure 241). The predicted polypeptide precursor is 600 amino acids long (Figure 242). The full-length PRO848 protein shown in Figure 242 has an estimated molecular weight of about 68,536 daltons. Clone DNA59839-1461
10 has been deposited with ATCC on June 16, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO848 polypeptide suggests that it may be a novel sialyltransferase. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced sequence identity between the PRO848 amino acid sequence and at least the following Dayhoff
15 sequences, P_R78619 (GalNAc-alpha-2, 6-sialyltransferase), CAAG5_CHICK (alpha-n-acetylgalactosamide alpha-2, 6-sialyltransferase), HSU14550_1, CAG6_HUMAN and P_R63217 (human alpha-2, 3-sialyltransferase).

EXAMPLE 107: Isolation of cDNA Clones Encoding Human PRO1097

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single
20 EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90)
25 or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56006.

In light of an observed sequence homology between the DNA56006 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2408105, the Incyte EST clone 2408105 was purchased
30 and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 243 and is herein designated as DNA59841-1460.

The entire nucleotide sequence of DNA59841-1460 is shown in Figure 243 (SEQ ID NO:348). Clone DNA59841-1460 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 3-5 and ending at the stop codon at nucleotide positions 276-278 of SEQ ID NO:348 (Figure 243).
35 The predicted polypeptide precursor is 91 amino acids long (Figure 244). The full-length PRO1097 protein shown in Figure 244 has an estimated molecular weight of about 10,542 daltons and a pI of about 10.04. Clone DNA59841-1460 has been deposited with ATCC on July 1, 1998. It is understood that the deposited clone has

the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analyzing Figure 244, the signal peptide is at about amino acids 1-20 of SEQ ID NO:349. The glycoprotease family protein domain starts at about amino acid 56, and the acyltransferase ChoActase/COT/CPT family peptide starts at about amino acid 49 of SEQ ID NO:349.

5

EXAMPLE 108: Isolation of cDNA clones Encoding Human PRO1153

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56008.

In light of an observed sequence homology between the DNA56008 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2472409, the Incyte EST clone 2472409 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 245 and is herein designated as DNA59842-1502.

The full length clone shown in Figure 245 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 92-94 and ending at the stop codon found at nucleotide positions 683-685 (Figure 245; SEQ ID NO:350). The predicted polypeptide precursor (Figure 246, SEQ ID NO:351) is 197 amino acids long. PRO1153 has a calculated molecular weight of approximately 21,540 daltons and an estimated pI of approximately 8.31. Clone DNA59842-1502 has been deposited with ATCC and is assigned ATCC deposit no. 209982. It is understood that the correct and actual sequence is in the deposited clone while herein are present representations based on current sequencing techniques which may have minor errors.

Based on a WU-BLAST2 sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO1153 shows some amino acid sequence identity to the following Dayhoff designations: S57447; SOYHRGPC_1; S46965; P_P82971; VCPHEROPH_1; EXTN_TOBAC; MLCB2548_9; ANXA_RABIT; JC5437 and SSGP_VOLCA.

EXAMPLE 109: Isolation of cDNA clones Encoding Human PRO1154

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in

Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56025.

In light of an observed sequence homology between the DNA56025 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2169375, the Incyte EST clone 2169375 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 247 and is herein designated as DNA59846-1503.

The full length clone shown in Figure 247 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 86-88 and ending at the stop codon found at nucleotide positions 2909-2911 (Figure 247; SEQ ID NO:352). The predicted polypeptide precursor (Figure 248, SEQ ID NO:353) is 941 amino acids long. PRO1154 has a calculated molecular weight of approximately 107,144 daltons and an estimated pI of approximately 6.26. Clone DNA59846-1503 has been deposited with ATCC and is assigned ATCC deposit no. 209978.

Based on a WU-BLAST2 sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO1154 shows sequence identity to at least the following Dayhoff designations: AB011097_1, AMPN_HUMAN, RNU76997_1, I59331, GEN14047, HSU62768_1, P_R51281, CET07F10_1, SSU66371_1, and AMPRE_HUMAN.

EXAMPLE 110: Isolation of cDNA clones Encoding Human PRO1181

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database, designated herein as 82468. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56029.

In light of an observed sequence homology between the DNA56029 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2186536, the Incyte EST clone 2186536 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 249 and is herein designated as DNA59847-1511.

Clone DNA59847-1511 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 17-19 and ending at the stop codon at nucleotide positions 1328-1330 (Figure 249). The predicted polypeptide precursor is 437 amino acids long (Figure 250). The full-length PRO1181 protein shown in Figure 250 has an estimated molecular weight of about 46,363 daltons and a pI of about 6.22. Analysis of the full-length PRO1181 sequence shown in Figure 250 (SEQ ID NO:355) evidences the presence of the

following: a signal peptide from about amino acid 1 to about amino acid 15, potential N-glycosylation sites from about amino acid 46 to about amino acid 49, from about amino acid 189 to about amino acid 192 and from about amino acid 382 to about amino acid 385 and amino acid sequence blocks having homology to Ly-6/u-PAR domain proteins from about amino acid 287 to about amino acid 300 and from about amino acid 98 to about amino acid 111. Clone DNA59847-1511 has been deposited with ATCC on August 4, 1998 and is assigned

5 ATCC deposit no. 203098.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 250 (SEQ ID NO:355), evidenced homology between the PRO1181 amino acid sequence and the following Dayhoff sequences: AF041083_1, P_W26579, RNMAGPIAN_1, CELT13C2_2, LMSAP2GN_1, S61882, CEF35C5_12, DP87_DICDI, GIU47631_1 and

10 P_R07092.

EXAMPLE 111: Isolation of cDNA clones Encoding Human PRO1182

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database, designated herein as 146647. This EST cluster sequence was

15 then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and

20 assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56033.

In light of an observed sequence homology between the DNA56033 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2595195, the Incyte EST clone 2595195 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.

25 The sequence of this cDNA insert is shown in Figure 251 and is herein designated as DNA59848-1512.

Clone DNA59848-1512 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 67-69 and ending at the stop codon at nucleotide positions 880-882 (Figure 251). The predicted polypeptide precursor is 271 amino acids long (Figure 252). The full-length PRO1182 protein shown in Figure 252 has an estimated molecular weight of about 28,665 daltons and a pI of about 5.33. Analysis of

30 the full-length PRO1182 sequence shown in Figure 252 (SEQ ID NO:357) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 25, an amino acid block having homology to C-type lectin domain proteins from about amino acid 247 to about amino acid 256 and an amino acid sequence block having homology to C1q domain proteins from about amino acid 44 to about amino acid 77. Clone DNA59848-1512 has been deposited with ATCC on August 4, 1998 and is assigned ATCC deposit

35 no. 203088.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 252 (SEQ ID NO:357), evidenced significant

homology between the PRO1182 amino acid sequence and the following Dayhoff sequences: PSPD_BOVIN, CL43_BOVIN, CONG_BOVIN, P_W18780, P_R45005, P_R53257 and CELEGAP7_1.

EXAMPLE 112: Isolation of cDNA clones Encoding Human PRO1155

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single
 5 EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of
 expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary
 EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The
 homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in
Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90)
 10 or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with
 the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence
 obtained therefrom is herein designated DNA56102.

In light of an observed sequence homology between the DNA56102 consensus sequence and an EST
 sequence encompassed within the Incyte EST clone no. 2858870, the Incyte EST clone 2858870 was purchased
 15 and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.
 The sequence of this cDNA insert is shown in Figure 253 and is herein designated as DNA59849-1504.

The full length clone shown in Figure 253 contained a single open reading frame with an apparent
 translational initiation site at nucleotide positions 158-160 and ending at the stop codon found at nucleotide
 positions 563-565 (Figure 253; SEQ ID NO:358). The predicted polypeptide precursor (Figure 254, SEQ ID
 20 NO:359) is 135 amino acids long. PRO1155 has a calculated molecular weight of approximately 14,833 daltons
 and an estimated pI of approximately 9.78. Clone DNA59849-1504 has been deposited with ATCC and is
 assigned ATCC deposit no. 209986. It is understood that the actual clone has the correct sequence whereas
 herein are only representations which are prone to minor sequencing errors.

Based on a WU-BLAST2 sequence alignment analysis (using the ALIGN computer program) of the full-
 25 length sequence, PRO1155 shows some amino acid sequence identity with the following Dayhoff designations:
 TKNK_BOVIN; PVB19X587_1; AF019049_1; P_W00948; S72864; P_W00949; I62742; AF038501_1;
 TKNG_HUMAN; and YAT1_RHOBL. Based on the information provided herein, PRO1155 may play a role
 in providing neuroprotection and cognitive enhancement.

30 EXAMPLE 113: Isolation of cDNA clones Encoding Human PRO1156

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single
 EST cluster sequence from the Incyte database, designated herein as 138851. This EST cluster sequence was
 then compared to a variety of expressed sequence tag (EST) databases which included public EST databases
 (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to
 35 identify existing homologies. The homology search was performed using the computer program BLAST or
 BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a
 BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and

assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56261.

In light of an observed sequence homology between the DNA56261 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3675191, the Incyte EST clone 3675191 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.

5 The sequence of this cDNA insert is shown in Figure 255 and is herein designated as DNA59853-1505.

The full length clone shown in Figure 255 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 212-214 and ending at the stop codon found at nucleotide positions 689-691 (Figure 255; SEQ ID NO:360). The predicted polypeptide precursor (Figure 256, SEQ ID NO:361) is 159 amino acids long. PRO1156 has a calculated molecular weight of approximately 17,476 daltons, 10 an estimated pI of approximately 9.15, a signal peptide sequence at about amino acids 1 to about 22, and potential N-glycosylation sites at about amino acids 27-30 and 41-44.

Clone DNA59853-1505 was deposited with the ATCC on June 16, 1998 and is assigned ATCC deposit no. 209985.

15 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis (using the ALIGN computer program) of the full-length sequence shown in Figure 256 (SEQ ID NO:361), revealed some homology between the PRO1156 amino acid sequence and the following Dayhoff sequences: D45027_1, P_R79914, JC5309, KBF2_HUMAN, AF010144_1, GEN14351, S68681, P_R79915, ZMTAC_3, and HUMCPGO_1.

20 EXAMPLE 114: Isolation of cDNA Clones Encoding Human PRO1098

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The 25 homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56377.

30 In light of an observed sequence homology between the DNA56377 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3050917, the Incyte EST clone 3050917 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 257 and is herein designated as DNA59854-1459.

The entire nucleotide sequence of DNA59854-1459 is shown in Figure 257 (SEQ ID NO:362). Clone 35 DNA59854-1459 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 58-60 and ending at the stop codon at nucleotide positions 292-294 of SEQ ID NO:362 (Figure 257). The predicted polypeptide precursor is 78 amino acids long (Figure 258). The full-length PRO1098 protein

shown in Figure 258 has an estimated molecular weight of about 8,396 daltons and a pI of about 7.66. Clone DNA59854-1459 has been deposited with ATCC on June 16, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

5 Analyzing Figure 258, a signal peptide appears to be at about amino acids 1-19 of SEQ ID NO:363, an N-glycosylation site appears to be at about amino acids 37-40 of SEQ ID NO:363, and N-myristoylation sites appear to be at about 15-20, 19-24 and 60-65 of SEQ ID NO:363.

EXAMPLE 115: Isolation of cDNA clones Encoding Human PRO1127

10 Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90)

15 or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57959.

In light of an observed sequence homology between the DNA57959 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 685126, the Merck EST clone 685126 was purchased

20 and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 259 and is herein designated as DNA60283-1484.

The full length clone shown in Figure 259 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 126-128 and ending at the stop codon found at nucleotide positions 327-329 (Figure 259; SEQ ID NO:364). The predicted polypeptide precursor (Figure 260, SEQ ID

25 NO:365) is 67 amino acids long including a signal peptide at about 1-29 of SEQ ID NO:365. PRO1127 has a calculated molecular weight of approximately 7,528 daltons and an estimated pI of approximately 4.95. Clone DNA60283-1484 was deposited with the ATCC on July 1, 1998 and is assigned ATCC deposit no. 203043. It is understood that the deposited clone has the actual sequence, whereas representations which may have minor sequencing errors are presented herein.

30 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 260 (SEQ ID NO:365), revealed some homology between the PRO1127 amino acid sequence and the following Dayhoff sequences: AF037218_48, P_W09638, HBA_HETPO, S39821, KR2_EBV, CET20D3_8, HCU37630_1, HS193B12_10, S40012 and TRITUBC_1.

35

EXAMPLE 116: Isolation of cDNA clones Encoding Human PRO1126

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56250.

In light of an observed sequence homology between the DNA56250 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1437250, the Incyte EST clone 1437250 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 261 and is herein designated as DNA60615-1483.

Clone DNA60615-1483 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 110-112 and ending at the stop codon at nucleotide positions 1316-1318 (Figure 261). The predicted polypeptide precursor is 402 amino acids long (Figure 262). The full-length PRO1126 protein shown in Figure 262 has an estimated molecular weight of about 45,921 daltons and a pI of about 8.60. Analysis of the full-length PRO1126 sequence shown in Figure 262 (SEQ ID NO:367) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 25 and potential N-glycosylation sites from about amino acid 66 to about amino acid 69, from about amino acid 138 to about amino acid 141 and from about amino acid 183 to about amino acid 186. Clone DNA60615-1483 has been deposited with ATCC on June 16, 1998 and is assigned ATCC deposit no. 209980.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 262 (SEQ ID NO:367), evidenced significant homology between the PRO1126 amino acid sequence and the following Dayhoff sequences: I73636, NOMR_HUMAN, MMUSMYOC3_1, HS454G6_1, P_R98225, RNU78105_1, RNU72487_1, AF035301_1, CEELC48E7_4 and CEF11C3_3.

EXAMPLE 117: Isolation of cDNA clones Encoding Human PRO1125

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence

obtained therefrom is herein designated DNA56540.

In light of an observed sequence homology between the DNA56540 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1486114, the Incyte EST clone 1486114 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 263 and is herein designated as DNA60615-1483.

- 5 The full length clone shown in Figure 263 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 47-49 and ending at the stop codon found at nucleotide positions 1388-1390 (Figure 263; SEQ ID NO:368). The predicted polypeptide precursor (Figure 264, SEQ ID NO:369) is 447 amino acids long. PRO1125 has a calculated molecular weight of approximately 49,798 daltons and an estimated pI of approximately 9.78. Clone DNA60619-1482 has been deposited with ATCC and is assigned
10 ATCC deposit no. 209993. It is understood that the clone has the actual sequence and that the sequences herein are representations based on current techniques which may be prone to minor errors.

- Based on a WU-BLAST2 sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO1125 shows some sequence identity with the following Dayhoff designations: RCO1_NEUCR; S58306; PKWA_THECU; S76086; P_R85881; HET1_PODAN; SPU92792_1;
15 APAF_HUMAN; S76414 and S59317.

EXAMPLE 118: Isolation of cDNA clones Encoding Human PRO1186

- Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of
20 expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with
25 the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56748.

- In light of an observed sequence homology between the DNA56748 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3476792, the Incyte EST clone 3476792 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.
30 The sequence of this cDNA insert is shown in Figure 265 and is herein designated as DNA60621-1516.

- The full length clone shown in Figure 265 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 91-93 and ending at the stop codon found at nucleotide positions 406-408 (Figure 265; SEQ ID NO:370). The predicted polypeptide precursor (Figure 266, SEQ ID NO:371) is 105 amino acids long. The signal peptide is at amino acids 1-19 of SEQ ID NO:371. PRO1186 has a
35 calculated molecular weight of approximately 11,715 daltons and an estimated pI of approximately 9.05. Clone DNA60621-1516 was deposited with the ATCC on August 4, 1998 and is assigned ATCC deposit no. 203091.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 266 (SEQ ID NO:371), revealed some sequence identity between the PRO1186 amino acid sequence and the following Dayhoff sequences: VPRA_DENPO, LFE4_CHICK, AF034208_1, AF030433_1, A55035, COL_RABIT, CELB0507_9, S67826_1, S34665 and CRU73817_1.

5

EXAMPLE 119: Isolation of cDNA clones Encoding Human PRO1198

An initial DNA sequence referred to herein as DNA52083 was identified using a yeast screen in a human umbilical vein endothelial cell cDNA library that preferentially represents the 5' ends of the primary cDNA clones. DNA52083 was compared to ESTs from public databases (e.g., GenBank), and a proprietary
10 EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA), using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)]. The ESTs were clustered and assembled into a consensus DNA sequence using the computer program "phrap" (Phil Green, University of Washington, Seattle, Washington). One or more of the ESTs was obtained from human breast skin tissue biopsy. This consensus sequence is designated herein as DNA52780.

15

In light of an observed sequence homology between the DNA52780 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3852910, the Incyte EST clone 3852910 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 267 and is herein designated as DNA60622-1525.

20

The full length DNA60622-1525 clone shown in Figure 267 (SEQ ID NO:372) contained a single open reading frame with an apparent translational initiation site at nucleotide positions 54 to 56 and ending at the stop codon found at nucleotide positions 741 to 743. The predicted polypeptide precursor, which is shown in Figure 268 (SEQ ID NO:373), is 229 amino acids long. PRO1198 has a calculated molecular weight of approximately 25,764 daltons and an estimated pI of approximately 9.17. There is a signal peptide sequence at about amino acids 1 through 34. There is sequence identity with glycosyl hydrolases family 31 protein at about amino acids
25 142 to about 175.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 268 (SEQ ID NO:373), revealed some homology between the PRO1198 amino acid sequence and the following Dayhoff sequences: ATF6H11_6, UCRI_RAT, TOBSUP2NT_1, RCUERF3_1, AMU88186_1, P_W22485, S56579, AF040711_1, DPP4_PIG.

30

Clone DNA60622-1525 was been deposited with the ATCC on August 4, 1998, and is assigned ATCC deposit no. 203090.

EXAMPLE 120: Isolation of cDNA clones Encoding Human PRO1158

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single
35 EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The

homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57248.

5 In light of an observed sequence homology between the DNA57248 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2640776, the Incyte EST clone 2640776 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 269 and is herein designated as DNA60625-1507.

10 The full length clone shown in Figure 269 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 163 to 165 and ending at the stop codon found at nucleotide positions 532 to 534 (Figure 269; SEQ ID NO:374). The predicted polypeptide precursor (Figure 270, SEQ ID NO:375) is 123 amino acids long. PRO1158 has a calculated molecular weight of approximately 13,113 daltons and an estimated pI of approximately 8.53. Additional features include a signal peptide sequence at about amino acids 1-19, a transmembrane domain at about amino acids 56-80, and a potential N-glycosylation site at
15 about amino acids 36-39. Clone DNA60625-1507 was deposited with the ATCC on June 16, 1998 and is assigned ATCC deposit no. 209975.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 270 (SEQ ID NO:375), revealed some homology between the PRO1158 amino acid sequence and the following Dayhoff sequences: ATAC00310510F18A8.10,
20 P_R85151, PHS2_SOLTU, RNMHCIBAC_1, RNA1FMHC_1, I68771, RNRT1A10G_1, PTPA_HUMAN, HUMGACA_1, and CHKPTPA_1.

EXAMPLE 121: Isolation of cDNA clones Encoding Human PRO1159

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single
25 EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90)
30 or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57221.

In light of an observed sequence homology between the DNA57221 consensus sequence and an EST
35 sequence encompassed within the Incyte EST clone no. 376776, the Incyte EST clone 376776 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 271 and is herein designated as DNA60627-1508.

Clone DNA60627-1508 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 92-94 and ending at the stop codon at nucleotide positions 362-364 (Figure 271). The predicted polypeptide precursor is 90 amino acids long (Figure 272). The full-length PRO1159 protein shown in Figure 272 has an estimated molecular weight of about 9,840 daltons and a pI of about 10.13. Analysis of the full-length PRO1159 sequence shown in Figure 272 (SEQ ID NO:377) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 15 and a potential N-glycosylation site from about amino acid 38 to about amino acid 41. Clone DNA60627-1508 has been deposited with ATCC on August 4, 1998 and is assigned ATCC deposit no. 203092.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 272 (SEQ ID NO:377), evidenced significant homology between the PRO1159 amino acid sequence and the following Dayhoff sequences: AF016494_6, AF036708_20, DSSCUTE_1, D89100_1, S28060, MEFA_XENLA, AF020798_12, G70065, E64423, JQ2005.

EXAMPLE 122: Isolation of cDNA clones Encoding Human PRO1124

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56035.

In light of an observed sequence homology between the DNA56035 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2767646, the Incyte EST clone 2767646 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 273 and is herein designated as DNA60629-1481.

The full length clone shown in Figure 273 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 25-27 and ending at the stop codon found at nucleotide positions 2782-2784 (Figure 273; SEQ ID NO:378). The predicted polypeptide precursor (Figure 274, SEQ ID NO:379) is 919 amino acids long. PRO1124 has a calculated molecular weight of approximately 101,282 daltons and an estimated pI of approximately 5.37. Clone DNA60629-1481 has been deposited with the ATCC and is assigned ATCC deposit no. 209979. It is understood that the deposited clone has the actual sequence, whereas only representations based on current sequencing techniques which may include normal and minor errors, are provided herein.

Based on a WU-BLAST2 sequence alignment analysis of the full-length sequence, PRO1124 shows significant amino acid sequence identity to a chloride channel protein and to ECAM-1. Specifically, the following Dayhoff designations were identified as having sequence identity with PRO1124: ECLC_BOVIN,

AF001261_1, P_W06548, SSC6A10_1, AF004355_1, S76691, AF017642, BYU06866_2, CSA_DICDI and SAU47139_2.

EXAMPLE 123: Isolation of cDNA clones Encoding Human PRO1287

An expressed sequence tag (EST) DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST was identified which showed homology to the fringe protein. This EST sequence was then compared to various EST databases including public EST databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify homologous EST sequences. The comparison was performed using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)]. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). This consensus sequence obtained is herein designated DNA40568.

Based on the DNA40568 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1287. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, *supra*, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CTCGGGGAAAGGGACTTGATGTTGG-3' (SEQ ID NO:382)

reverse PCR primer 1 5'-GCGAAGGTGAGCCTCTATCTCGTGCC-3' (SEQ ID NO:383)

reverse PCR primer 2 5'-CAGCCTACACGTATTGAGG-3' (SEQ ID NO:384)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA40568 sequence which had the following nucleotide sequence

hybridization probe

5'-CAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCC-3' (SEQ ID NO:385).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO1287 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human bone marrow tissue. The cDNA libraries used to isolated the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD;

pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1287 (designated herein as DNA61755-1554 [Figure 275, SEQ ID NO:380]) and the derived protein sequence for PRO1287.

- 5 The entire nucleotide sequence of DNA61755-1554 is shown in Figure 275 (SEQ ID NO:380). The full length clone contained a single open reading frame with an apparent translational initiation site at nucleotide positions 655-657 and a stop signal at nucleotide positions 2251-2253 (Figure 275, SEQ ID NO:380). The predicted polypeptide precursor is 532 amino acids long, has a calculated molecular weight of approximately 61,351 daltons and an estimated pI of approximately 8.77. Analysis of the full-length PRO1287 sequence shown
- 10 in Figure 276 (SEQ ID NO:381) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 27 and potential N-glycosylation sites from about amino acid 315 to about amino acid 318 and from about amino acid 324 to about amino acid 327. Clone DNA61755-1554 has been deposited with ATCC on August 11, 1998 and is assigned ATCC deposit no. 203112.

- An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence
- 15 alignment analysis of the full-length sequence shown in Figure 276 (SEQ ID NO:381), evidenced significant homology between the PRO1287 amino acid sequence and the following Dayhoff sequences: CET24D1_1, EZRI_BOVIN, GGU19889_1, CC3_YEAST, S74244, NALS_MOUSE, MOES_PIG, S28660, S44860 and YNA4_CAEEL.

20 EXAMPLE 124: Isolation of cDNA clones Encoding Human PRO1312

DNA55773 was identified in a human fetal kidney cDNA library using a yeast screen that preferentially represents the 5' ends of the primary cDNA clones. Based on the DNA55773 sequence, oligonucleotides were synthesized for use as probes to isolate a clone of the full-length coding sequence for PRO1312.

- The full length DNA61873-1574 clone shown in Figure 277 (SEQ ID NO:386) contained a single open
- 25 reading frame with an apparent translational initiation site at nucleotide positions 7-9 and ending at the stop codon found at nucleotide positions 643-645. The predicted polypeptide precursor is 212 amino acids long (Figure 278, SEQ ID NO:387). PRO1312 has a calculated molecular weight of approximately 24,024 daltons and an estimated pI of approximately 6.26. Other features include a signal peptide at about amino acids 1-14; a transmembrane domain at about amino acids 141-160, and potential N-glycosylation sites at about amino acids
- 30 76-79 and 93-96.

- An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 278 (SEQ ID NO:387), revealed some homology between the PRO1312 amino acid sequence and the following Dayhoff sequences: GCINTALPH_1, GIBMUC1A_1, P_R96298, AF001406_1, PVU88874_1, P_R85151, AF041409_1, CELC50F2_7, C45875,
- 35 and AB009510_21.

Clone DNA61873-1574 has been deposited with ATCC and is assigned ATCC deposit no. 203132.

EXAMPLE 125: Isolation of cDNA clones Encoding Human PRO1192

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein DNA35924. Based on the DNA35924 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1192.

PCR primers (forward and reverse) were synthesized:

forward PCR primer: 5'-CCGAGGCCATCTAGAGGCCAGAGC-3' (SEQ ID NO:390)

reverse PCR primer: 5'-ACAGGCAGAGCCAATGGCCAGAGC-3' (SEQ ID NO:391).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35924 sequence which had the following nucleotide sequence:

hybridization probe:

5'-GAGAGGACTGCGGGAGTTTGGGACCTTTGTGCAGACGTGCTCATG-3' (SEQ ID NO:392).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1192 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal liver and spleen tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1192 designated herein as DNA62814-1521 and shown in Figure 279 (SEQ ID NO:388); and the derived protein sequence for PRO1192 which is shown in Figure 280 (SEQ ID NO:389).

The entire coding sequence of PRO1192 is shown in Figure 279 (SEQ ID NO:388). Clone DNA62814-1521 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 121-123 and an apparent stop codon at nucleotide positions 766-768. The predicted polypeptide precursor is 215 amino acids long. The predicted polypeptide precursor has the following features: a signal peptide at about amino acids 1-21; a transmembrane domain at about amino acids 153-176; potential N-glycosylation sites at about amino acids 39-42 and 118-121; and homology with myelin P0 proteins at about amino acids 27-68 and 99-128 of Figure 280. The full-length PRO1192 protein shown in Figure 280 has an estimated molecular weight of about 24,484 daltons and a pI of about 6.98.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 280 (SEQ ID NO:389), revealed homology between the PRO1192 amino acid sequence and the following Dayhoff sequences: GEN12838, MYP0_HUMAN, AF049498_1, GEN14531, P_W14146, HS46KDA_1, CINB_RAT, OX2G_RAT, D87018_1, and D86996_2.

Clone DNA62814-1521 was deposited with the ATCC on August 4, 1998, and is assigned ATCC deposit no. 203093.

EXAMPLE 126: Isolation of cDNA clones Encoding Human PRO1160

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA40650. Based on the DNA40650 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1160.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GCTCCCTGATCTTCATGTCACCACC-3' (SEQ ID NO:395)

reverse PCR primer 5'-CAGGGACACACTCTACCATTCGGGAG-3' (SEQ ID NO:396)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA40650 sequence which had the following nucleotide sequence

hybridization probe

5'-CCATCTTTCTGGTCTCTGCCCAGAATCCGACAACAGCTGCTC-3' (SEQ ID NO:397)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1160 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human breast tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1160 (designated herein as DNA62872-1509 [Figure 281, SEQ ID NO: 393]) and the derived protein sequence for PRO1160.

The entire nucleotide sequence of DNA62872-1509 is shown in Figure 281 (SEQ ID NO:393). Clone DNA62872-1509 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 40-42 and ending at the stop codon at nucleotide positions 310-312 (Figure 281). The predicted polypeptide precursor is 90 amino acids long (Figure 282). The full-length PRO1160 protein shown in Figure 282 has an estimated molecular weight of about 9,039 daltons and a pI of about 4.37. Analysis of the full-length PRO1160 sequence shown in Figure 282 (SEQ ID NO:394) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 19 and a protein kinase C phosphorylation site from about amino acid 68 to about amino acid 70. Clone DNA62872-1509 has been deposited with ATCC on August 4, 1998 and is assigned ATCC deposit no. 203100.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 282 (SEQ ID NO:394), evidenced significant homology between the PRO1160 amino acid sequence and the following Dayhoff sequences: B30305, GEN13490, I53641, S53363, HA34_BRELC, SP96_DICDI, S36326, SSU51197_10, MUC1_XENLA, TCU32448_1 and AF000409_1.

EXAMPLE 127: Isolation of cDNA clones Encoding Human PRO1187

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of

expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57726.

In light of an observed sequence homology between the DNA57726 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 358563, the Incyte EST clone 358563 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 283 and is herein designated as DNA62876-1517.

The full length clone shown in Figure 283 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 121-123 and ending at the stop codon found at nucleotide positions 481-483 (Figure 283; SEQ ID NO:398). The predicted polypeptide precursor (Figure 284, SEQ ID NO:399) is 120 amino acids long. The signal peptide is at about amino acids 1-17 of SEQ ID NO:399. PRO1187 has a calculated molecular weight of approximately 12,925 daltons and an estimated pI of approximately 9.46. Clone DNA62876-1517 was deposited with the ATCC on August 4, 1998 and is assigned ATCC deposit no. 203095. It is understood that the deposited clone contains the actual sequence and that the representations herein may have minor sequencing errors.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 284 (SEQ ID NO:399), revealed some sequence identity (and therefore some relation) between the PRO1187 amino acid sequence and the following Dayhoff sequences: MGNENDOBX_1, CELF41G3_9, AMPG_STRLI, HSBBOVHERL_2, LEEXTEN10_1, AF029958_1 and P_W04957.

EXAMPLE 128: Isolation of cDNA clones Encoding Human PRO1185

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56426.

In light of an observed sequence homology between the DNA56426 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3284411, the Incyte EST clone 3284411 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.

The sequence of this cDNA insert is shown in Figure 285 and is herein designated as DNA62881-1515.

The full length DNA62881-1515 clone shown in Figure 285 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 4-6 and ending at the stop codon found at nucleotide positions 598-600 (Figure 285; SEQ ID NO:400). The predicted polypeptide precursor (Figure 286, SEQ ID NO:401) is 198 amino acids long. The signal peptide is at about amino acids 1-21 of SEQ ID NO:401.

5 PRO1185 has a calculated molecular weight of approximately 22,105 daltons and an estimated pI of approximately 7.73. Clone DNA62881-1515 has been deposited with the ATCC and is assigned ATCC deposit no. 203096.

10 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 286 (SEQ ID NO:401), revealed some sequence identity between the PRO1185 amino acid sequence and the following Dayhoff sequences: TUP1_YEAST, AF041382_1, MAOM_SOLTU, SPPBPHU9_1, I41024, EPCPLCFail_1, HSPLEC_1, YKL4_CAEEL, A44643, TGU65922_1.

EXAMPLE 129: Isolation of cDNA clones Encoding Human PRO1345

15 A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA47364. Based on the DNA47364 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1345.

20 PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCTGGTTATCCCCAGGAAGTCCGAC-3' (SEQ ID NO:404)

reverse PCR primer 5'-CTCTTGCTGCTGCGACAGGCCTC-3' (SEQ ID NO:405)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA47364 sequence which had the following nucleotide sequence

25 hybridization probe

5'-CGCCCTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGAC-3' (SEQ ID NO:406)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1345 gene using the probe oligonucleotide and one of the PCR primers. RNA
30 for construction of the cDNA libraries was isolated from human breast carcinoma tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1345 (designated herein as DNA64852-1589 [Figure 287, SEQ ID NO:402]) and the derived protein sequence for PRO1345.

35 The entire nucleotide sequence of DNA64852-1589 is shown in Figure 287 (SEQ ID NO:402). Clone DNA64852-1589 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 7-9 or 34-36 and ending at the stop codon at nucleotide positions 625-627 (Figure 287). The predicted polypeptide precursor is 206 amino acids long (Figure 288). The full-length PRO1345 protein shown in Figure

288 has an estimated molecular weight of about 23,190 daltons and a pI of about 9.40. Analysis of the full-length PRO1345 sequence shown in Figure 288 (SEQ ID NO:403) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 31 or from about amino acid 10 to about amino acid 31 and a C-type lectin domain signature sequence from about amino acid 176 to about amino acid 190. Clone DNA64852-1589 has been deposited with ATCC on August 18, 1998 and is assigned ATCC deposit no. 203127.

5 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 288 (SEQ ID NO:403), evidenced significant homology between the PRO1345 amino acid sequence and the following Dayhoff sequences: BTU22298_1, TETN_CARSP, TETN_HUMAN, MABA_RAT, S34198, P_W13144, MACMBPA_1, A46274, PSPD_RAT AND P_R32188.

10 EXAMPLE 130: Isolation of cDNA clones Encoding Human PRO1245

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary
15 EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence
20 obtained therefrom is herein designated DNA56019.

In light of an observed sequence homology between the DNA56019 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1327836, the Incyte EST clone 1327836 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 289 and is herein designated as DNA64884-1527.

25 The full length clone shown in Figure 289 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 79-81 and ending at the stop codon found at nucleotide positions 391-393 (Figure 289; SEQ ID NO:407). The predicted polypeptide precursor (Figure 290, SEQ ID NO:408) is 104 amino acids long, with a signal peptide sequence at about amino acid 1 to about amino acid 18. PRO1245 has a calculated molecular weight of approximately 10,100 daltons and an estimated pI of
30 approximately 8.76.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 290 (SEQ ID NO:408), revealed some homology between the PRO1245 amino acid sequence and the following Dayhoff sequences: SYA_THETH, GEN11167, MTV044_4, AB011151_1, RLAJ2750_3, SNEIPTRA_1, S63624, C28391, A37907, and S14064.

35 Clone DNA64884-1245 was deposited with the ATCC on August 25, 1998 and is assigned ATCC deposit no. 203155.

EXAMPLE 131: Isolation of cDNA clones Encoding Human PRO1358

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

In light of an observed sequence homology between the consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 88718, the Incyte EST clone 88718 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 291 and is herein designated as DNA64890-1612.

The full length clone shown in Figure 291 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 86 through 88 and ending at the stop codon found at nucleotide positions 1418 through 1420 (Figure 291; SEQ ID NO:409). The predicted polypeptide precursor (Figure 292, SEQ ID NO:410) is 444 amino acids long. The signal peptide is at about amino acids 1-18 of SEQ ID NO:410. PRO1358 has a calculated molecular weight of approximately 50,719 daltons and an estimated pI of approximately 8.82. Clone DNA64890-1612 was deposited with the ATCC on August 18, 1998 and is assigned ATCC deposit no. 203131.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 292 (SEQ ID NO:410), revealed sequence identity between the PRO1358 amino acid sequence and the following Dayhoff sequences: P_W07607, AB000545_1, AB000546_1, A1AT_RAT, AB015164_1, P_P50021, COTR_CAVPO, and HAMHPP_1. The variants claimed in this application exclude these sequences.

EXAMPLE 132: Isolation of cDNA clones Encoding Human PRO1195

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55716.

In light of an observed sequence homology between the DNA55716 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3252980, the Incyte EST clone 3252980 was purchased

and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 293 and is herein designated as DNA65412-1523.

The full length clone shown in Figure 293 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 58-60 and ending at the stop codon found at nucleotide positions 511-513 (Figure 293; SEQ ID NO:411). The predicted polypeptide precursor (Figure 294, SEQ ID NO:412) is 151 amino acids long. The signal sequence is at about amino acids 1-22 of SEQ ID NO:412. PRO1195 has a calculated molecular weight of approximately 17,277 daltons and an estimated pI of approximately 5.33. Clone DNA65412-1523 was deposited with the ATCC on August 4, 1998 and is assigned ATCC deposit no. 203094.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 294 (SEQ ID NO:412), revealed some sequence identity between the PRO1195 amino acid sequence and the following Dayhoff sequences: MMU28486_1, AF044205_1, P_W31186, CELK03C7_1, F69034, EF1A_METVA, AF024540_1, SSU90353_1, MRSP_STAAU and P_R97680.

EXAMPLE 133: Isolation of cDNA clones Encoding Human PRO1270

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57951.

In light of an observed sequence homology between the DNA57951 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 124878, the Merck EST clone 124878 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 295 and is herein designated as DNA66308-1537.

Clone DNA66308-1537 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 103-105 and ending at the stop codon at nucleotide positions 1042-1044 (Figure 295). The predicted polypeptide precursor is 313 amino acids long (Figure 296). The full-length PRO1270 protein shown in Figure 296 has an estimated molecular weight of about 34,978 daltons and a pI of about 5.71. Analysis of the full-length PRO1270 sequence shown in Figure 296 (SEQ ID NO:414) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 16, a potential N-glycosylation site from about amino acid 163 to about amino acid 166 and glycosaminoglycan attachment sites from about amino acid 74 to about amino acid 77 and from about amino acid 289 to about amino acid 292. Clone DNA66308-1537 has been deposited with ATCC on August 25, 1998 and is assigned ATCC deposit no. 203159.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 296 (SEQ ID NO:414), evidenced significant homology between the PRO1270 amino acid sequence and the following Dayhoff sequences: XLU86699_1, S49589, FIBA_PARPA, FIBB_HUMAN, P_R47189, AF004326_1, DRTENASCN_1, AF004327_1, P_W01411 and FIBG_BOVIN.

5

EXAMPLE 134: Isolation of cDNA clones Encoding Human PRO1271

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57955.

15

In light of an observed sequence homology between the DNA57955 consensus sequence and an EST sequence encompassed within the Merck EST clone no. AA625350, the Merck EST clone AA625350 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 297 and is herein designated as DNA66309-1538.

20

Clone DNA66309-1538 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 94-96 and ending at the stop codon at nucleotide positions 718-720 (Figure 297). The predicted polypeptide precursor is 208 amino acids long (Figure 298). The full-length PRO1271 protein shown in Figure 298 has an estimated molecular weight of about 21,531 daltons and a pI of about 8.99. Analysis of the full-length PRO1271 sequence shown in Figure 298 (SEQ ID NO:416) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 31 and a transmembrane domain from about amino acid 166 to about amino acid 187. Clone DNA66309-1538 has been deposited with ATCC on September 15, 1998 and is assigned ATCC deposit no. 203235.

25

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 298 (SEQ ID NO:416), evidenced significant homology between the PRO1271 amino acid sequence and the following Dayhoff sequences: S57180, S63257, AGA1_YEAST, BPU43599_1, YS8A_CAEEL, S67570, LSU54556_2, S70305, VGLX_HSVEB, and D88733_1.

30

EXAMPLE 135: Isolation of cDNA clones Encoding Human PRO1375

35

A Merck/Wash. U. database was searched and a Merck EST was identified. This sequence was then put in a program which aligns it with other sequences from the Swiss-Prot public database, public EST databases (e.g., GenBank, Merck/Wash. U.), and a proprietary EST database (LIFESEQ®, Incyte

Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)] as a comparison of the extracellular domain (ECD) protein sequences to a 6 frame translation of the EST sequences. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

A consensus DNA sequence was assembled relative to other EST sequences using phrap. This consensus sequence is designated herein "DNA67003".

Based on the DNA67003 consensus sequence, the nucleic acid (SEQ ID NO:417) was identified in a human pancreas library. DNA sequencing of the clone gave the full-length DNA sequence for PRO1375 and the derived protein sequence for PRO1375.

The entire coding sequence of PRO1375 is shown in Figure 299 (SEQ ID NO:417). Clone DNA67004-1614 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 104-106 and an apparent stop codon at nucleotide positions 698-700 of SEQ ID NO:417. The predicted polypeptide precursor is 198 amino acids long. The transmembrane domains are at about amino acids 11-28 (type II) and 103-125 of SEQ ID NO:418. Clone DNA67004-1614 has been deposited with ATCC and is assigned ATCC deposit no. 203115. The full-length PRO1375 protein shown in Figure 300 has an estimated molecular weight of about 22,531 daltons and a pI of about 8.47.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 300 (SEQ ID NO:418), revealed sequence identity between the PRO1375 amino acid sequence and the following Dayhoff sequences: AF026198_5, CELR12C12_5, S73465, Y011_MYCPN, S64538_1, P_P8150, MUVSHPO10_1, VSH_MUMPL and CVU59751_5.

EXAMPLE 136: Isolation of cDNA clones Encoding Human PRO1385

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57952.

In light of an observed sequence homology between the DNA57952 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3129630, the Incyte EST clone 3129630 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 301 and is herein designated as DNA68869-1610.

Clone DNA68869-1610 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 26-28 and ending at the stop codon at nucleotide positions 410-412 (Figure 301). The predicted polypeptide precursor is 128 amino acids long (Figure 302). The full-length PRO1385 protein shown in Figure 302 has an estimated molecular weight of about 13,663 daltons and a pI of about 10.97. Analysis of the full-length PRO1385 sequence shown in Figure 302 (SEQ ID NO:420) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 28, and glycosylaminoglycan attachment sites from about amino acid 82 to about amino acid 85 and from about amino acid 91 to about amino acid 94. Clone DNA68869-1610 has been deposited with ATCC on August 25, 1998 and is assigned ATCC deposit no. 203164.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 302 (SEQ ID NO:420), evidenced low homology between the PRO1385 amino acid sequence and the following Dayhoff sequences: CELT14A8_1, LMNACHRA1_1, HXD9_HUMAN, CHKCMLF_1, HS5PP34_2, DMDRING_1, A37107_1, MMLUNGENE_1, PUM_DROME and DMU25117_1.

EXAMPLE 137: Isolation of cDNA clones Encoding Human PRO1387

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56259.

In light of an observed sequence homology between the DNA56259 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3507924, the Incyte EST clone 3507924 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 303 and is herein designated as DNA68872-1620.

Clone DNA68872-1620 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 85-87 and ending at the stop codon at nucleotide positions 1267-1269 (Figure 303). The predicted polypeptide precursor is 394 amino acids long (Figure 304). The full-length PRO1387 protein shown in Figure 304 has an estimated molecular weight of about 44,339 daltons and a pI of about 7.10. Analysis of the full-length PRO1387 sequence shown in Figure 304 (SEQ ID NO:422) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 19, a transmembrane domain from about amino acid 275 to about amino acid 296, potential N-glycosylation sites from about amino acid 76 to about amino acid 79, from about amino acid 231 to about amino acid 234, from about amino acid 302 to about amino acid 305, from about amino acid 307 to about amino acid 310 and from about amino acid 376 to about amino acid

379, and amino acid sequence blocks having homology to myelin p0 protein from about amino acid 210 to about amino acid 239 and from about amino acid 92 to about amino acid 121. Clone DNA68872-1620 has been deposited with ATCC on August 25, 1998 and is assigned ATCC deposit no. 203160.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 304 (SEQ ID NO:422), evidenced significant homology between the PRO1387 amino acid sequence and the following Dayhoff sequences: P_W36955, MYPO_HETFR, HS46KDA_1, AF049498_1, MYOO_HUMAN, AF030454_1, A53268, SHPTCRA_1, P_W14146 and GEN12838.

EXAMPLE 138: Isolation of cDNA clones Encoding Human PRO1384

10 A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA54192. Based on the DNA54192 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1384.

PCR primers (forward and reverse) were synthesized:

15 forward PCR primer 5'-TGCAGCCCCTGTGACACAACTGG-3' (SEQ ID NO:425)

reverse PCR primer 5'-CTGAGATAACCGAGCCATCCTCCCAC-3' (SEQ ID NO:426)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA54192 sequence which had the following nucleotide sequence:

hybridization probe

20 5'-GGAGATAGCTGCTATGGGTTCTTCAGGCACAACTTAACATGGGAAG-3' (SEQ ID NO:427)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1384 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal liver.

25 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1384 (designated herein as DNA71159-1617 [Figure 305, SEQ ID NO:423]; and the derived protein sequence for PRO1384.

The entire coding sequence of PRO1384 is shown in Figure 305 (SEQ ID NO:423). Clone DNA71159-1617 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 182-184 and an apparent stop codon at nucleotide positions 869-871. The predicted polypeptide precursor is 229 amino acids long. The full-length PRO1384 protein shown in Figure 306 has an estimated molecular weight of about 26,650 daltons and a pI of about 8.76. Additional features include a type II transmembrane domain at about amino acids 32-57, and potential N-glycosylation sites at about amino acids 68-71, 120-123, and 134-137.

35 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 306 (SEQ ID NO:424), revealed homology between the PRO1384 amino acid sequence and the following Dayhoff sequences: AF054819_1, HSAJ1687_1, AF009511_1, AB010710_1, GEN13595, HSAJ673_1, GEN13961, AB005900_1, LECH_CHICK, AF021349_1,

and NK13_RAT.

Clone DNA71159-1617 has been deposited with ATCC and is assigned ATCC deposit no. 203135.

EXAMPLE 139: Use of PRO as a hybridization probe

The following method describes use of a nucleotide sequence encoding PRO as a hybridization probe.

5 DNA comprising the coding sequence of full-length or mature PRO as disclosed herein is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO-derived probe to the filters is performed in a
10 solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO can then be identified using standard techniques known in the art.

15

EXAMPLE 140: Expression of PRO in *E. coli*

This example illustrates preparation of an unglycosylated form of PRO by recombinant expression in *E. coli*.

The DNA sequence encoding PRO is initially amplified using selected PCR primers. The primers
20 should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode
25 for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant
30 colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

35 After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO protein can then be purified using a metal chelating column under conditions that allow tight binding of the

protein.

PRO may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110 *fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq)*). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D. 600 of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g (NH₄)₂SO₄, 0.71 g sodium citrate•2H₂O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM MgSO₄) and grown for approximately 20-30 hours at 30°C with shaking. Samples are removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

E. coli paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentrifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The clarified extract is loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4°C for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin

from the samples.

Fractions containing the desired folded PRO polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

5 Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 141: Expression of PRO in mammalian cells

This example illustrates preparation of a potentially glycosylated form of PRO by recombinant expression in mammalian cells.

10 The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO DNA using ligation methods such as described in Sambrook et al., supra. The resulting vector is called pRK5-PRO.

15 In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 μ g pRK5-PRO DNA is mixed with about 1 μ g DNA encoding the VA RNA gene [Thimmappaya et al., Cell, 31:543 (1982)] and dissolved in 500 μ l of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl_2 . To this mixture is added, dropwise, 500 μ l of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO_4 , and a precipitate is allowed to form for 10 minutes at 25°C. The
20 precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μ Ci/ml ^{35}S -cysteine and 200 μ Ci/ml ^{35}S -methionine. After
25 a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO may be introduced into 293 cells transiently using the dextran sulfate
30 method described by Sompayrac et al., Proc. Natl. Acad. Sci., 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μ g pRK5-PRO DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μ g/ml bovine insulin and
35 0.1 μ g/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO can be expressed in CHO cells. The pRK5-PRO can be transfected into CHO cells using known reagents such as CaPO₄ or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ³⁵S-methionine. After determining the presence of PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO can then be concentrated and purified by any selected method.

Epitope-tagged PRO may also be expressed in host CHO cells. The PRO may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO can then be concentrated and purified by any selected method, such as by Ni²⁺-chelate affinity chromatography.

PRO may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas et al., Nucl. Acids Res. 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect[®] (Quiagen), Dosper[®] or Fugene[®] (Boehringer Mannheim). The cells are grown as described in Lucas et al., supra. Approximately 3 x 10⁷ cells are frozen in an ampule for further growth and production as described below.

The ampules containing the plasmid DNA are thawed by placement into water bath and mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 mL of selective media (0.2 μ m filtered PS20 with 5% 0.2 μ m diafiltered fetal bovine serum). The cells are then aliquoted into a 100 mL spinner containing 90 mL of selective media. After 1-2 days, the cells are transferred into a 250 mL spinner filled with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, 250 mL, 500 mL and 2000 mL spinners are seeded with 3 x 10⁵ cells/mL. The cell media is exchanged with fresh media by

centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Patent No. 5,122,469, issued June 16, 1992 may actually be used. A 3L production spinner is seeded at 1.2×10^6 cells/mL. On day 0, the cell number pH is determined. On day 1, the spinner is sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, the temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability dropped below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22 μ m filter. The filtrate was either stored at 4°C or immediately loaded onto columns for purification.

For the poly-His tagged constructs, the proteins are purified using a Ni-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The conditioned medium is pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275 μ L of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

25 EXAMPLE 142: Expression of PRO in Yeast

The following method describes recombinant expression of PRO in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO from the ADH2/GAPDH promoter. DNA encoding PRO and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of PRO. For secretion, DNA encoding PRO can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native PRO signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO may further be purified using selected column chromatography resins.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

5 EXAMPLE 143: Expression of PRO in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO in Baculovirus-infected insect cells.

The sequence coding for PRO is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding PRO or the desired portion of the coding sequence of PRO such as the sequence encoding the extracellular domain of a transmembrane protein or the sequence encoding the mature protein if the protein is extracellular is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

15 Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (PharMingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as described by O'Reilley et al., Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

20 Expressed poly-his tagged PRO can then be purified, for example, by Ni²⁺-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl₂; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni²⁺-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A₂₈₀ with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A₂₈₀ baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or Western blot with Ni²⁺-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His₁₀-tagged PRO are pooled and dialyzed against loading buffer.

35 Alternatively, purification of the IgG tagged (or Fc tagged) PRO can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 144: Preparation of Antibodies that Bind PRO

This example illustrates preparation of monoclonal antibodies which can specifically bind PRO.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, *supra*. Immunogens that may be employed include purified PRO, fusion proteins containing PRO, and cells expressing recombinant PRO on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against PRO. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against PRO is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

EXAMPLE 145: Purification of PRO Polypeptides Using Specific Antibodies

Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody

is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

Such an immunoaffinity column is utilized in the purification of PRO polypeptide by preparing a fraction from cells containing PRO polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO polypeptide (*e.g.*, high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO polypeptide binding (*e.g.*, a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and PRO polypeptide is collected.

EXAMPLE 146: Drug Screening

This invention is particularly useful for screening compounds by using PRO polypeptides or binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with an PRO polypeptide or fragment thereof and assaying (i) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or fragment is typically labeled. After suitable incubation, free PRO polypeptide or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO polypeptide or to interfere with the PRO polypeptide/cell complex.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the

solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding PRO polypeptide specifically compete with a test compound for binding to PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO polypeptide.

5

EXAMPLE 147: Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptide of interest (*i.e.*, a PRO polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide *in vivo* (*c.f.*, Hodgson, Bio/Technology, 9: 19-21 (1991)).

In one approach, the three-dimensional structure of the PRO polypeptide, or of an PRO polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, Biochemistry, 31:7796-7801 (1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, J. Biochem., 113:742-746 (1993).

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

35

Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209, USA (ATCC):

Table 2

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
5	DNA16422-1209	209929	June 2, 1998
	DNA16435-1208	209930	June 2, 1998
	DNA21624-1391	209917	June 2, 1998
	DNA23334-1392	209918	June 2, 1998
10	DNA26288-1239	209792	April 21, 1998
	DNA26843-1389	203099	August 4, 1998
	DNA26844-1394	209926	June 2, 1998
	DNA30862-1396	209920	June 2, 1998
	DNA35680-1212	209790	April 21, 1998
15	DNA40621-1440	209922	June 2, 1998
	DNA44161-1434	209907	May 27, 1998
	DNA44694-1500	203114	August 11, 1998
	DNA45495-1550	203156	August 25, 1998
	DNA47361-1154	209431	November 7, 1997
20	DNA47394-1572	203109	August 11, 1998
	DNA48320-1433	209904	May 27, 1998
	DNA48334-1435	209924	June 2, 1998
	DNA48606-1479	203040	July 1, 1998
	DNA49141-1431	203003	June 23, 1998
25	DNA49142-1430	203002	June 23, 1998
	DNA49143-1429	203013	June 23, 1998
	DNA49647-1398	209919	June 2, 1998
	DNA49819-1439	209931	June 2, 1998
	DNA49820-1427	209932	June 2, 1998
30	DNA49821-1562	209981	June 16, 1998
	DNA52192-1369	203042	July 1, 1998
	DNA52598-1518	203107	August 11, 1998
	DNA53913-1490	203162	August 25, 1998
	DNA53978-1443	209983	June 16, 1998
35	DNA53996-1442	209921	June 2, 1998
	DNA56041-1416	203012	June 23, 1998
	DNA56047-1456	209948	June 9, 1998
	DNA56050-1455	203011	June 23, 1998
	DNA56110-1437	203113	August 11, 1998
40	DNA56113-1378	203049	July 1, 1998
	DNA56410-1414	209923	June 2, 1998
	DNA56436-1448	209902	May 27, 1998
	DNA56855-1447	203004	June 23, 1998
	DNA56859-1445	203019	June 23, 1998
45	DNA56860-1510	209952	June 9, 1998
	DNA56865-1491	203022	June 23, 1998
	DNA56866-1342	203023	June 23, 1998
	DNA56868-1209	203024	June 23, 1998
	DNA56869-1545	203161	August 25, 1998
50	DNA56870-1492	209925	June 2, 1998
	DNA57033-1403	209905	May 27, 1998
	DNA57037-1444	209903	May 27, 1998
	DNA57129-1413	209977	June 16, 1998

	DNA57690-1374	209950	June 9, 1998
	DNA57693-1424	203008	June 23, 1998
	DNA57694-1341	203017	June 23, 1998
	DNA57695-1340	203006	June 23, 1998
	DNA57699-1412	203020	June 23, 1998
5	DNA57702-1476	209951	June 9, 1998
	DNA57704-1452	209953	June 9, 1998
	DNA57708-1411	203021	June 23, 1998
	DNA57710-1451	203048	July 1, 1998
	DNA57711-1501	203047	July 1, 1998
10	DNA57827-1493	203045	July 1, 1998
	DNA57834-1339	209954	June 9, 1998
	DNA57836-1338	203025	June 23, 1998
	DNA57838-1337	203014	June 23, 1998
	DNA57844-1410	203010	June 23, 1998
15	DNA58721-1475	203110	August 11, 1998
	DNA58723-1588	203133	August 18, 1998
	DNA58737-1473	203136	August 18, 1998
	DNA58743-1609	203154	August 25, 1998
	DNA58846-1409	209957	June 9, 1998
20	DNA58848-1472	209955	June 9, 1998
	DNA58849-1494	209958	June 9, 1998
	DNA58850-1495	209956	June 9, 1998
	DNA58853-1423	203016	June 23, 1998
	DNA58855-1422	203018	June 23, 1998
25	DNA59205-1421	203009	June 23, 1998
	DNA59211-1450	209960	June 9, 1998
	DNA59213-1487	209959	June 9, 1998
	DNA59214-1449	203046	July 1, 1998
	DNA59215-1425	209961	June 9, 1998
30	DNA59220-1514	209962	June 9, 1998
	DNA59488-1603	203157	August 25, 1998
	DNA59493-1420	203050	July 1, 1998
	DNA59497-1496	209941	June 4, 1998
	DNA59588-1571	203106	August 11, 1998
35	DNA59603-1419	209944	June 9, 1998
	DNA59605-1418	203005	June 23, 1998
	DNA59606-1471	209945	June 9, 1998
	DNA59607-1497	209957	June 9, 1998
	DNA59609-1470	209963	June 9, 1998
40	DNA59610-1559	209990	June 16, 1998
	DNA59612-1466	209947	June 9, 1998
	DNA59613-1417	203007	June 23, 1998
	DNA59616-1465	209991	June 16, 1998
	DNA59619-1464	203041	July 1, 1998
45	DNA59620-1463	209989	June 16, 1998
	DNA59625-1498	209992	June 17, 1998
	DNA59767-1489	203108	August 11, 1998
	DNA59776-1600	203128	August 18, 1998
	DNA59777-1480	203111	August 11, 1998
50	DNA59820-1549	203129	August 18, 1998
	DNA59827-1426	203089	August 4, 1998
	DNA59828-1608	203158	August 25, 1998
	DNA59838-1462	209976	June 16, 1998
	DNA59839-1461	209988	June 16, 1998
55	DNA59841-1460	203044	July 1, 1998
	DNA59842-1502	209982	June 16, 1998

	DNA59846-1503	209978	June 16, 1998
	DNA59847-1511	203098	August 4, 1998
	DNA59848-1512	203088	August 4, 1998
	DNA59849-1504	209986	June 16, 1998
	DNA59853-1505	209985	June 16, 1998
5	DNA59854-1459	209974	June 16, 1998
	DNA60283-1484	203043	July 1, 1998
	DNA60615-1483	209980	June 16, 1998
	DNA60619-1482	209993	June 16, 1998
	DNA60621-1516	203091	August 4, 1998
10	DNA60622-1525	203090	August 4, 1998
	DNA60625-1507	209975	June 16, 1998
	DNA60627-1508	203092	August 4, 1998
	DNA60629-1481	209979	June 16, 1998
	DNA61755-1554	203112	August 11, 1998
15	DNA61873-1574	203132	August 18, 1998
	DNA62814-1521	203093	August 4, 1998
	DNA62872-1509	203100	August 4, 1998
	DNA62876-1517	203095	August 4, 1998
	DNA62881-1515	203096	August 4, 1998
20	DNA64852-1589	203127	August 18, 1998
	DNA64884-1527	203155	August 25, 1998
	DNA64890-1612	203131	August 18, 1998
	DNA65412-1523	203094	August 4, 1998
	DNA66308-1537	203159	August 25, 1998
25	DNA66309-1538	203235	September 15, 1998
	DNA67004-1614	203115	August 11, 1998
	DNA68869-1610	203164	August 25, 1998
	DNA68872-1620	203160	August 25, 1998
30	DNA71159-1617	203135	August 18, 1998

These deposit were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to USC §122 and the Commissioner's rules pursuant thereto (including 37 CFR §1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the

deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition
5 to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. Isolated nucleic acid having at least 80% sequence identity to a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:6), Figure 6 (SEQ ID NO:8), Figure 9 (SEQ ID NO:14), Figure 12 (SEQ ID NO:20), Figure 15 (SEQ ID NO:23), Figure 18 (SEQ ID NO:28), Figure 20 (SEQ ID NO:30), Figure 23 (SEQ ID NO:33), Figure 25 (SEQ ID NO:36), Figure 27 (SEQ ID NO:41), Figure 30 (SEQ ID NO:47), Figure 32 (SEQ ID NO:52), Figure 34 (SEQ ID NO:57), Figure 36 (SEQ ID NO:62), Figure 38 (SEQ ID NO:67), Figure 41 (SEQ ID NO:73), Figure 47 (SEQ ID NO:84), Figure 49 (SEQ ID NO:95), Figure 51 (SEQ ID NO:97), Figure 53 (SEQ ID NO:99), Figure 57 (SEQ ID NO:103), Figure 64 (SEQ ID NO:113), Figure 66 (SEQ ID NO:115), Figure 68 (SEQ ID NO:117), Figure 70 (SEQ ID NO:119), Figure 72 (SEQ ID NO:124), Figure 74 (SEQ ID NO:129), Figure 76 (SEQ ID NO:135), Figure 79 (SEQ ID NO:138), Figure 83 (SEQ ID NO:146), Figure 85 (SEQ ID NO:148), Figure 88 (SEQ ID NO:151), Figure 90 (SEQ ID NO:153), Figure 93 (SEQ ID NO:156), Figure 95 (SEQ ID NO:158), Figure 97 (SEQ ID NO:160), Figure 99 (SEQ ID NO:165), Figure 101 (SEQ ID NO:167), Figure 103 (SEQ ID NO:169), Figure 105 (SEQ ID NO:171), Figure 109 (SEQ ID NO:175), Figure 111 (SEQ ID NO:177), Figure 113 (SEQ ID NO:179), Figure 115 (SEQ ID NO:181), Figure 117 (SEQ ID NO:183), Figure 120 (SEQ ID NO:189), Figure 122 (SEQ ID NO:194), Figure 125 (SEQ ID NO:197), Figure 127 (SEQ ID NO:199), Figure 129 (SEQ ID NO:201), Figure 131 (SEQ ID NO:203), Figure 133 (SEQ ID NO:205), Figure 135 (SEQ ID NO:207), Figure 137 (SEQ ID NO:209), Figure 139 (SEQ ID NO:211), Figure 141 (SEQ ID NO:213), Figure 144 (SEQ ID NO:216), Figure 147 (SEQ ID NO:219), Figure 149 (SEQ ID NO:221), Figure 151 (SEQ ID NO:223), Figure 153 (SEQ ID NO:225), Figure 155 (SEQ ID NO:227), Figure 157 (SEQ ID NO:229), Figure 159 (SEQ ID NO:231), Figure 161 (SEQ ID NO:236), Figure 163 (SEQ ID NO:241), Figure 165 (SEQ ID NO:246), Figure 167 (SEQ ID NO:248), Figure 169 (SEQ ID NO:250), Figure 171 (SEQ ID NO:253), Figure 174 (SEQ ID NO:256), Figure 176 (SEQ ID NO:258), Figure 178 (SEQ ID NO:260), Figure 180 (SEQ ID NO:262), Figure 182 (SEQ ID NO:264), Figure 184 (SEQ ID NO:266), Figure 186 (SEQ ID NO:268), Figure 188 (SEQ ID NO:270), Figure 190 (SEQ ID NO:272), Figure 192 (SEQ ID NO:274), Figure 194 (SEQ ID NO:276), Figure 196 (SEQ ID NO:278), Figure 198 (SEQ ID NO:281), Figure 200 (SEQ ID NO:283), Figure 202 (SEQ ID NO:285), Figure 204 (SEQ ID NO:287), Figure 206 (SEQ ID NO:289), Figure 208 (SEQ ID NO:291), Figure 210 (SEQ ID NO:293), Figure 212 (SEQ ID NO:295), Figure 214 (SEQ ID NO:297), Figure 216 (SEQ ID NO:299), Figure 218 (SEQ ID NO:301), Figure 220 (SEQ ID NO:303), Figure 226 (SEQ ID NO:309), Figure 228 (SEQ ID NO:314), Figure 230 (SEQ ID NO:319), Figure 233 (SEQ ID NO:326), Figure 235 (SEQ ID NO:334), Figure 238 (SEQ ID NO:340), Figure 240 (SEQ ID NO:345), Figure 242 (SEQ ID NO:347), Figure 244 (SEQ ID NO:349), Figure 246 (SEQ ID NO:351), Figure 248 (SEQ ID NO:353), Figure 250 (SEQ ID NO:355), Figure 252 (SEQ ID NO:357), Figure 254 (SEQ ID NO:359), Figure 256 (SEQ ID NO:361), Figure 258 (SEQ ID NO:363), Figure 260 (SEQ ID NO:365), Figure 262 (SEQ ID NO:367), Figure 264 (SEQ ID NO:369), Figure 266 (SEQ ID NO:371), Figure 268 (SEQ ID NO:373), Figure 270 (SEQ ID NO:375), Figure 272 (SEQ ID NO:377), Figure 274 (SEQ ID NO:379), Figure 276 (SEQ ID NO:381), Figure 278 (SEQ ID NO:387), Figure 280 (SEQ ID NO:389), Figure 282 (SEQ ID NO:394), Figure 284 (SEQ ID NO:399), Figure 286 (SEQ

ID NO:401), Figure 288 (SEQ ID NO:403), Figure 290 (SEQ ID NO:408), Figure 292 (SEQ ID NO:410), Figure 294 (SEQ ID NO:412), Figure 296 (SEQ ID NO:414), Figure 298 (SEQ ID NO:416), Figure 300 (SEQ ID NO:418), Figure 302 (SEQ ID NO:420), Figure 304 (SEQ ID NO:422) and Figure 306 (SEQ ID NO:424).

2. The nucleic acid sequence of Claim 1, wherein said nucleotide sequence comprises a nucleotide
5 sequence selected from the group consisting of the sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:5), Figure 5 (SEQ ID NO:7), Figure 8 (SEQ ID NO:13), Figure 11 (SEQ ID NO:19), Figure 14 (SEQ ID NO:22), Figure 17 (SEQ ID NO:27), Figure 19 (SEQ ID NO:29), Figure 22 (SEQ ID NO:32), Figure 24 (SEQ ID NO:35), Figure 26 (SEQ ID NO:40), Figure 29 (SEQ ID NO:46), Figure 31 (SEQ ID NO:51), Figure 33 (SEQ ID NO:56), Figure 35 (SEQ ID NO:61), Figure 37 (SEQ ID NO:66), Figure 40 (SEQ ID NO:72),
10 Figure 46 (SEQ ID NO:83), Figure 48 (SEQ ID NO:94), Figure 50 (SEQ ID NO:96), Figure 52 (SEQ ID NO:98), Figure 56 (SEQ ID NO:102), Figure 63 (SEQ ID NO:112), Figure 65 (SEQ ID NO:114), Figure 67 (SEQ ID NO:116), Figure 69 (SEQ ID NO:118), Figure 71 (SEQ ID NO:123), Figure 73 (SEQ ID NO:128), Figure 75 (SEQ ID NO:134), Figure 78 (SEQ ID NO:137), Figure 82 (SEQ ID NO:145), Figure 84 (SEQ ID NO:147), Figure 87 (SEQ ID NO:150), Figure 89 (SEQ ID NO:152), Figure 92 (SEQ ID NO:155), Figure 94
15 (SEQ ID NO:157), Figure 96 (SEQ ID NO:159), Figure 98 (SEQ ID NO:164), Figure 100 (SEQ ID NO:166), Figure 102 (SEQ ID NO:168), Figure 104 (SEQ ID NO:170), Figure 108 (SEQ ID NO:174), Figure 110 (SEQ ID NO:176), Figure 112 (SEQ ID NO:178), Figure 114 (SEQ ID NO:180), Figure 116 (SEQ ID NO:182), Figure 119 (SEQ ID NO:188), Figure 121 (SEQ ID NO:193), Figure 124 (SEQ ID NO:196), Figure 126 (SEQ ID NO:198), Figure 128 (SEQ ID NO:200), Figure 130 (SEQ ID NO:202), Figure 132 (SEQ ID NO:204),
20 Figure 134 (SEQ ID NO:206), Figure 136 (SEQ ID NO:208), Figure 138 (SEQ ID NO:210), Figure 140 (SEQ ID NO:212), Figure 143 (SEQ ID NO:215), Figure 146 (SEQ ID NO:218), Figure 148 (SEQ ID NO:220), Figure 150 (SEQ ID NO:222), Figure 152 (SEQ ID NO:224), Figure 154 (SEQ ID NO:226), Figure 156 (SEQ ID NO:228), Figure 158 (SEQ ID NO:230), Figure 160 (SEQ ID NO:235), Figure 162 (SEQ ID NO:240), Figure 164 (SEQ ID NO:245), Figure 166 (SEQ ID NO:247), Figure 168 (SEQ ID NO:249), Figure 170 (SEQ ID NO:252), Figure 173 (SEQ ID NO:255), Figure 175 (SEQ ID NO:257), Figure 177 (SEQ ID NO:259),
25 Figure 179 (SEQ ID NO:261), Figure 181 (SEQ ID NO:263), Figure 183 (SEQ ID NO:265), Figure 185 (SEQ ID NO:267), Figure 187 (SEQ ID NO:269), Figure 189 (SEQ ID NO:271), Figure 191 (SEQ ID NO:273), Figure 193 (SEQ ID NO:275), Figure 195 (SEQ ID NO:277), Figure 197 (SEQ ID NO:280), Figure 199 (SEQ ID NO:282), Figure 201 (SEQ ID NO:284), Figure 203 (SEQ ID NO:286), Figure 205 (SEQ ID NO:288),
30 Figure 207 (SEQ ID NO:290), Figure 209 (SEQ ID NO:292), Figure 211 (SEQ ID NO:294), Figure 213 (SEQ ID NO:296), Figure 215 (SEQ ID NO:298), Figure 217 (SEQ ID NO:300), Figure 219 (SEQ ID NO:302), Figure 225 (SEQ ID NO:308), Figure 227 (SEQ ID NO:313), Figure 229 (SEQ ID NO:318), Figure 232 (SEQ ID NO:325), Figure 234 (SEQ ID NO:333), Figure 237 (SEQ ID NO:339), Figure 239 (SEQ ID NO:344), Figure 241 (SEQ ID NO:346), Figure 243 (SEQ ID NO:348), Figure 245 (SEQ ID NO:350), Figure 247 (SEQ ID NO:352), Figure 249 (SEQ ID NO:354), Figure 251 (SEQ ID NO:356), Figure 253 (SEQ ID NO:358),
35 Figure 255 (SEQ ID NO:360), Figure 257 (SEQ ID NO:362), Figure 259 (SEQ ID NO:364), Figure 261 (SEQ ID NO:366), Figure 263 (SEQ ID NO:368), Figure 265 (SEQ ID NO:370), Figure 267 (SEQ ID NO:372),

Figure 269 (SEQ ID NO:374), Figure 271 (SEQ ID NO:376), Figure 273 (SEQ ID NO:378), Figure 275 (SEQ ID NO:380), Figure 277 (SEQ ID NO:386), Figure 279 (SEQ ID NO:388), Figure 281 (SEQ ID NO:393), Figure 283 (SEQ ID NO:398), Figure 285 (SEQ ID NO:400), Figure 287 (SEQ ID NO:402), Figure 289 (SEQ ID NO:407), Figure 291 (SEQ ID NO:409), Figure 293 (SEQ ID NO:411), Figure 295 (SEQ ID NO:413), Figure 297 (SEQ ID NO:415), Figure 299 (SEQ ID NO:417), Figure 301 (SEQ ID NO:419), Figure 303 (SEQ ID NO:421) and Figure 305 (SEQ ID NO:423).

3. The nucleic acid of Claim 1, wherein said nucleotide sequence comprises a nucleotide sequence selected from the group consisting of the full-length coding sequence of the sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:5), Figure 5 (SEQ ID NO:7), Figure 8 (SEQ ID NO:13), Figure 11 (SEQ ID NO:19), Figure 14 (SEQ ID NO:22), Figure 17 (SEQ ID NO:27), Figure 19 (SEQ ID NO:29), Figure 22 (SEQ ID NO:32), Figure 24 (SEQ ID NO:35), Figure 26 (SEQ ID NO:40), Figure 29 (SEQ ID NO:46), Figure 31 (SEQ ID NO:51), Figure 33 (SEQ ID NO:56), Figure 35 (SEQ ID NO:61), Figure 37 (SEQ ID NO:66), Figure 40 (SEQ ID NO:72), Figure 46 (SEQ ID NO:83), Figure 48 (SEQ ID NO:94), Figure 50 (SEQ ID NO:96), Figure 52 (SEQ ID NO:98), Figure 56 (SEQ ID NO:102), Figure 63 (SEQ ID NO:112), Figure 65 (SEQ ID NO:114), Figure 67 (SEQ ID NO:116), Figure 69 (SEQ ID NO:118), Figure 71 (SEQ ID NO:123), Figure 73 (SEQ ID NO:128), Figure 75 (SEQ ID NO:134), Figure 78 (SEQ ID NO:137), Figure 82 (SEQ ID NO:145), Figure 84 (SEQ ID NO:147), Figure 87 (SEQ ID NO:150), Figure 89 (SEQ ID NO:152), Figure 92 (SEQ ID NO:155), Figure 94 (SEQ ID NO:157), Figure 96 (SEQ ID NO:159), Figure 98 (SEQ ID NO:164), Figure 100 (SEQ ID NO:166), Figure 102 (SEQ ID NO:168), Figure 104 (SEQ ID NO:170), Figure 108 (SEQ ID NO:174), Figure 110 (SEQ ID NO:176), Figure 112 (SEQ ID NO:178), Figure 114 (SEQ ID NO:180), Figure 116 (SEQ ID NO:182), Figure 119 (SEQ ID NO:188), Figure 121 (SEQ ID NO:193), Figure 124 (SEQ ID NO:196), Figure 126 (SEQ ID NO:198), Figure 128 (SEQ ID NO:200), Figure 130 (SEQ ID NO:202), Figure 132 (SEQ ID NO:204), Figure 134 (SEQ ID NO:206), Figure 136 (SEQ ID NO:208), Figure 138 (SEQ ID NO:210), Figure 140 (SEQ ID NO:212), Figure 143 (SEQ ID NO:215), Figure 146 (SEQ ID NO:218), Figure 148 (SEQ ID NO:220), Figure 150 (SEQ ID NO:222), Figure 152 (SEQ ID NO:224), Figure 154 (SEQ ID NO:226), Figure 156 (SEQ ID NO:228), Figure 158 (SEQ ID NO:230), Figure 160 (SEQ ID NO:235), Figure 162 (SEQ ID NO:240), Figure 164 (SEQ ID NO:245), Figure 166 (SEQ ID NO:247), Figure 168 (SEQ ID NO:249), Figure 170 (SEQ ID NO:252), Figure 173 (SEQ ID NO:255), Figure 175 (SEQ ID NO:257), Figure 177 (SEQ ID NO:259), Figure 179 (SEQ ID NO:261), Figure 181 (SEQ ID NO:263), Figure 183 (SEQ ID NO:265), Figure 185 (SEQ ID NO:267), Figure 187 (SEQ ID NO:269), Figure 189 (SEQ ID NO:271), Figure 191 (SEQ ID NO:273), Figure 193 (SEQ ID NO:275), Figure 195 (SEQ ID NO:277), Figure 197 (SEQ ID NO:280), Figure 199 (SEQ ID NO:282), Figure 201 (SEQ ID NO:284), Figure 203 (SEQ ID NO:286), Figure 205 (SEQ ID NO:288), Figure 207 (SEQ ID NO:290), Figure 209 (SEQ ID NO:292), Figure 211 (SEQ ID NO:294), Figure 213 (SEQ ID NO:296), Figure 215 (SEQ ID NO:298), Figure 217 (SEQ ID NO:300), Figure 219 (SEQ ID NO:302), Figure 225 (SEQ ID NO:308), Figure 227 (SEQ ID NO:313), Figure 229 (SEQ ID NO:318), Figure 232 (SEQ ID NO:325), Figure 234 (SEQ ID NO:333), Figure 237 (SEQ ID NO:339), Figure 239 (SEQ ID NO:344), Figure 241 (SEQ ID NO:346), Figure 243 (SEQ ID NO:348), Figure 245 (SEQ ID

NO:350), Figure 247 (SEQ ID NO:352), Figure 249 (SEQ ID NO:354), Figure 251 (SEQ ID NO:356), Figure 253 (SEQ ID NO:358), Figure 255 (SEQ ID NO:360), Figure 257 (SEQ ID NO:362), Figure 259 (SEQ ID NO:364), Figure 261 (SEQ ID NO:366), Figure 263 (SEQ ID NO:368), Figure 265 (SEQ ID NO:370), Figure 267 (SEQ ID NO:372), Figure 269 (SEQ ID NO:374), Figure 271 (SEQ ID NO:376), Figure 273 (SEQ ID NO:378), Figure 275 (SEQ ID NO:380), Figure 277 (SEQ ID NO:386), Figure 279 (SEQ ID NO:388), Figure 281 (SEQ ID NO:393), Figure 283 (SEQ ID NO:398), Figure 285 (SEQ ID NO:400), Figure 287 (SEQ ID NO:402), Figure 289 (SEQ ID NO:407), Figure 291 (SEQ ID NO:409), Figure 293 (SEQ ID NO:411), Figure 295 (SEQ ID NO:413), Figure 297 (SEQ ID NO:415), Figure 299 (SEQ ID NO:417), Figure 301 (SEQ ID NO:419), Figure 303 (SEQ ID NO:421) or Figure 305 (SEQ ID NO:423).

10 4. Isolated nucleic acid which comprises the full-length coding sequence of the DNA deposited under any ATCC accession number shown in Table 2.

5 5. A vector comprising the nucleic acid of Claim 1.

15 6. The vector of Claim 5 operably linked to control sequences recognized by a host cell transformed with the vector.

 7. A host cell comprising the vector of Claim 5.

20 8. The host cell of Claim 7 wherein said cell is a CHO cell.

 9. The host cell of Claim 7 wherein said cell is an *E. coli*.

 10. The host cell of Claim 7 wherein said cell is a yeast cell.

25 11. A process for producing a PRO polypeptides comprising culturing the host cell of Claim 7 under conditions suitable for expression of said PRO polypeptide and recovering said PRO polypeptide from the cell culture.

30 12. Isolated PRO polypeptide having at least 80% sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:6), Figure 6 (SEQ ID NO:8), Figure 9 (SEQ ID NO:14), Figure 12 (SEQ ID NO:20), Figure 15 (SEQ ID NO:23), Figure 18 (SEQ ID NO:28), Figure 20 (SEQ ID NO:30), Figure 23 (SEQ ID NO:33), Figure 25 (SEQ ID NO:36), Figure 27 (SEQ ID NO:41), Figure 30 (SEQ ID NO:47), Figure 32 (SEQ ID NO:52), Figure 34 (SEQ ID NO:57), Figure 36 (SEQ ID NO:62), Figure 38 (SEQ ID NO:67), Figure 41 (SEQ ID NO:73), Figure 47 (SEQ ID NO:84), Figure 49 (SEQ ID NO:95), Figure 51 (SEQ ID NO:97), Figure 53 (SEQ ID NO:99), Figure 57 (SEQ ID NO:103), Figure 64 (SEQ ID NO:113), Figure 66 (SEQ ID NO:115), Figure 68

35

(SEQ ID NO:117), Figure 70 (SEQ ID NO:119), Figure 72 (SEQ ID NO:124), Figure 74 (SEQ ID NO:129), Figure 76 (SEQ ID NO:135), Figure 79 (SEQ ID NO:138), Figure 83 (SEQ ID NO:146), Figure 85 (SEQ ID NO:148), Figure 88 (SEQ ID NO:151), Figure 90 (SEQ ID NO:153), Figure 93 (SEQ ID NO:156), Figure 95 (SEQ ID NO:158), Figure 97 (SEQ ID NO:160), Figure 99 (SEQ ID NO:165), Figure 101 (SEQ ID NO:167), Figure 103 (SEQ ID NO:169), Figure 105 (SEQ ID NO:171), Figure 109 (SEQ ID NO:175), Figure 111 (SEQ ID NO:177), Figure 113 (SEQ ID NO:179), Figure 115 (SEQ ID NO:181), Figure 117 (SEQ ID NO:183), Figure 120 (SEQ ID NO:189), Figure 122 (SEQ ID NO:194), Figure 125 (SEQ ID NO:197), Figure 127 (SEQ ID NO:199), Figure 129 (SEQ ID NO:201), Figure 131 (SEQ ID NO:203), Figure 133 (SEQ ID NO:205), Figure 135 (SEQ ID NO:207), Figure 137 (SEQ ID NO:209), Figure 139 (SEQ ID NO:211), Figure 141 (SEQ ID NO:213), Figure 144 (SEQ ID NO:216), Figure 147 (SEQ ID NO:219), Figure 149 (SEQ ID NO:221), Figure 151 (SEQ ID NO:223), Figure 153 (SEQ ID NO:225), Figure 155 (SEQ ID NO:227), Figure 157 (SEQ ID NO:229), Figure 159 (SEQ ID NO:231), Figure 161 (SEQ ID NO:236), Figure 163 (SEQ ID NO:241), Figure 165 (SEQ ID NO:246), Figure 167 (SEQ ID NO:248), Figure 169 (SEQ ID NO:250), Figure 171 (SEQ ID NO:253), Figure 174 (SEQ ID NO:256), Figure 176 (SEQ ID NO:258), Figure 178 (SEQ ID NO:260), Figure 180 (SEQ ID NO:262), Figure 182 (SEQ ID NO:264), Figure 184 (SEQ ID NO:266), Figure 186 (SEQ ID NO:268), Figure 188 (SEQ ID NO:270), Figure 190 (SEQ ID NO:272), Figure 192 (SEQ ID NO:274), Figure 194 (SEQ ID NO:276), Figure 196 (SEQ ID NO:278), Figure 198 (SEQ ID NO:281), Figure 200 (SEQ ID NO:283), Figure 202 (SEQ ID NO:285), Figure 204 (SEQ ID NO:287), Figure 206 (SEQ ID NO:289), Figure 208 (SEQ ID NO:291), Figure 210 (SEQ ID NO:293), Figure 212 (SEQ ID NO:295), Figure 214 (SEQ ID NO:297), Figure 216 (SEQ ID NO:299), Figure 218 (SEQ ID NO:301), Figure 220 (SEQ ID NO:303), Figure 226 (SEQ ID NO:309), Figure 228 (SEQ ID NO:314), Figure 230 (SEQ ID NO:319), Figure 233 (SEQ ID NO:326), Figure 235 (SEQ ID NO:334), Figure 238 (SEQ ID NO:340), Figure 240 (SEQ ID NO:345), Figure 242 (SEQ ID NO:347), Figure 244 (SEQ ID NO:349), Figure 246 (SEQ ID NO:351), Figure 248 (SEQ ID NO:353), Figure 250 (SEQ ID NO:355), Figure 252 (SEQ ID NO:357), Figure 254 (SEQ ID NO:359), Figure 256 (SEQ ID NO:361), Figure 258 (SEQ ID NO:363), Figure 260 (SEQ ID NO:365), Figure 262 (SEQ ID NO:367), Figure 264 (SEQ ID NO:369), Figure 266 (SEQ ID NO:371), Figure 268 (SEQ ID NO:373), Figure 270 (SEQ ID NO:375), Figure 272 (SEQ ID NO:377), Figure 274 (SEQ ID NO:379), Figure 276 (SEQ ID NO:381), Figure 278 (SEQ ID NO:387), Figure 280 (SEQ ID NO:389), Figure 282 (SEQ ID NO:394), Figure 284 (SEQ ID NO:399), Figure 286 (SEQ ID NO:401), Figure 288 (SEQ ID NO:403), Figure 290 (SEQ ID NO:408), Figure 292 (SEQ ID NO:410), Figure 294 (SEQ ID NO:412), Figure 296 (SEQ ID NO:414), Figure 298 (SEQ ID NO:416), Figure 300 (SEQ ID NO:418), Figure 302 (SEQ ID NO:420), Figure 304 (SEQ ID NO:422) and Figure 306 (SEQ ID NO:424).

13. Isolated PRO polypeptide having at least 80% sequence identity to the amino acid sequence encoded by a nucleic acid molecule deposited under any ATCC accession number shown in Table 2.

14. A chimeric molecule comprising a polypeptide according to Claim 12 fused to a heterologous amino acid sequence.

15. The chimeric molecule of Claim 14 wherein said heterologous amino acid sequence is an epitope tag sequence.

16. The chimeric molecule of Claim 14 wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

5

17. An antibody which specifically binds to a PRO polypeptide according to Claim 12.

18. The antibody of Claim 17 wherein said antibody is a monoclonal antibody.

10

19. The antibody of Claim 17 wherein said antibody is a humanized antibody.

20. The antibody of Claim 17 wherein said antibody is an antibody fragment.

21. An isolated nucleic acid molecule which has at least 80% sequence identity to a nucleic acid
15 which comprises a nucleotide sequence selected from the group consisting of that shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:5), Figure 5 (SEQ ID NO:7), Figure 8 (SEQ ID NO:13), Figure 11 (SEQ ID NO:19), Figure 14 (SEQ ID NO:22), Figure 17 (SEQ ID NO:27), Figure 19 (SEQ ID NO:29), Figure 22 (SEQ ID NO:32), Figure 24 (SEQ ID NO:35), Figure 26 (SEQ ID NO:40), Figure 29 (SEQ ID NO:46), Figure 31 (SEQ ID NO:51), Figure 33 (SEQ ID NO:56), Figure 35 (SEQ ID NO:61), Figure 37 (SEQ ID NO:66), Figure
20 40 (SEQ ID NO:72), Figure 46 (SEQ ID NO:83), Figure 48 (SEQ ID NO:94), Figure 50 (SEQ ID NO:96), Figure 52 (SEQ ID NO:98), Figure 56 (SEQ ID NO:102), Figure 63 (SEQ ID NO:112), Figure 65 (SEQ ID NO:114), Figure 67 (SEQ ID NO:116), Figure 69 (SEQ ID NO:118), Figure 71 (SEQ ID NO:123), Figure 73 (SEQ ID NO:128), Figure 75 (SEQ ID NO:134), Figure 78 (SEQ ID NO:137), Figure 82 (SEQ ID NO:145), Figure 84 (SEQ ID NO:147), Figure 87 (SEQ ID NO:150), Figure 89 (SEQ ID NO:152), Figure 92 (SEQ ID
25 NO:155), Figure 94 (SEQ ID NO:157), Figure 96 (SEQ ID NO:159), Figure 98 (SEQ ID NO:164), Figure 100 (SEQ ID NO:166), Figure 102 (SEQ ID NO:168), Figure 104 (SEQ ID NO:170), Figure 108 (SEQ ID NO:174), Figure 110 (SEQ ID NO:176), Figure 112 (SEQ ID NO:178), Figure 114 (SEQ ID NO:180), Figure 116 (SEQ ID NO:182), Figure 119 (SEQ ID NO:188), Figure 121 (SEQ ID NO:193), Figure 124 (SEQ ID NO:196), Figure 126 (SEQ ID NO:198), Figure 128 (SEQ ID NO:200), Figure 130 (SEQ ID NO:202), Figure
30 132 (SEQ ID NO:204), Figure 134 (SEQ ID NO:206), Figure 136 (SEQ ID NO:208), Figure 138 (SEQ ID NO:210), Figure 140 (SEQ ID NO:212), Figure 143 (SEQ ID NO:215), Figure 146 (SEQ ID NO:218), Figure 148 (SEQ ID NO:220), Figure 150 (SEQ ID NO:222), Figure 152 (SEQ ID NO:224), Figure 154 (SEQ ID NO:226), Figure 156 (SEQ ID NO:228), Figure 158 (SEQ ID NO:230), Figure 160 (SEQ ID NO:235), Figure 162 (SEQ ID NO:240), Figure 164 (SEQ ID NO:245), Figure 166 (SEQ ID NO:247), Figure 168 (SEQ ID
35 NO:249), Figure 170 (SEQ ID NO:252), Figure 173 (SEQ ID NO:255), Figure 175 (SEQ ID NO:257), Figure 177 (SEQ ID NO:259), Figure 179 (SEQ ID NO:261), Figure 181 (SEQ ID NO:263), Figure 183 (SEQ ID NO:265), Figure 185 (SEQ ID NO:267), Figure 187 (SEQ ID NO:269), Figure 189 (SEQ ID NO:271), Figure

191 (SEQ ID NO:273), Figure 193 (SEQ ID NO:275), Figure 195 (SEQ ID NO:277), Figure 197 (SEQ ID NO:280), Figure 199 (SEQ ID NO:282), Figure 201 (SEQ ID NO:284), Figure 203 (SEQ ID NO:286), Figure 205 (SEQ ID NO:288), Figure 207 (SEQ ID NO:290), Figure 209 (SEQ ID NO:292), Figure 211 (SEQ ID NO:294), Figure 213 (SEQ ID NO:296), Figure 215 (SEQ ID NO:298), Figure 217 (SEQ ID NO:300), Figure 219 (SEQ ID NO:302), Figure 225 (SEQ ID NO:308), Figure 227 (SEQ ID NO:313), Figure 229 (SEQ ID NO:318), Figure 232 (SEQ ID NO:325), Figure 234 (SEQ ID NO:333), Figure 237 (SEQ ID NO:339), Figure 239 (SEQ ID NO:344), Figure 241 (SEQ ID NO:346), Figure 243 (SEQ ID NO:348), Figure 245 (SEQ ID NO:350), Figure 247 (SEQ ID NO:352), Figure 249 (SEQ ID NO:354), Figure 251 (SEQ ID NO:356), Figure 253 (SEQ ID NO:358), Figure 255 (SEQ ID NO:360), Figure 257 (SEQ ID NO:362), Figure 259 (SEQ ID NO:364), Figure 261 (SEQ ID NO:366), Figure 263 (SEQ ID NO:368), Figure 265 (SEQ ID NO:370), Figure 267 (SEQ ID NO:372), Figure 269 (SEQ ID NO:374), Figure 271 (SEQ ID NO:376), Figure 273 (SEQ ID NO:378), Figure 275 (SEQ ID NO:380), Figure 277 (SEQ ID NO:386), Figure 279 (SEQ ID NO:388), Figure 281 (SEQ ID NO:393), Figure 283 (SEQ ID NO:398), Figure 285 (SEQ ID NO:400), Figure 287 (SEQ ID NO:402), Figure 289 (SEQ ID NO:407), Figure 291 (SEQ ID NO:409), Figure 293 (SEQ ID NO:411), Figure 295 (SEQ ID NO:413), Figure 297 (SEQ ID NO:415), Figure 299 (SEQ ID NO:417), Figure 301 (SEQ ID NO:419), Figure 303 (SEQ ID NO:421) and Figure 305 (SEQ ID NO:423).

22. An isolated nucleic acid molecule which has at least 80% sequence identity to the full-length coding sequence of a nucleotide sequence selected from the group consisting of that shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:5), Figure 5 (SEQ ID NO:7), Figure 8 (SEQ ID NO:13), Figure 11 (SEQ ID NO:19), Figure 14 (SEQ ID NO:22), Figure 17 (SEQ ID NO:27), Figure 19 (SEQ ID NO:29), Figure 22 (SEQ ID NO:32), Figure 24 (SEQ ID NO:35), Figure 26 (SEQ ID NO:40), Figure 29 (SEQ ID NO:46), Figure 31 (SEQ ID NO:51), Figure 33 (SEQ ID NO:56), Figure 35 (SEQ ID NO:61), Figure 37 (SEQ ID NO:66), Figure 40 (SEQ ID NO:72), Figure 46 (SEQ ID NO:83), Figure 48 (SEQ ID NO:94), Figure 50 (SEQ ID NO:96), Figure 52 (SEQ ID NO:98), Figure 56 (SEQ ID NO:102), Figure 63 (SEQ ID NO:112), Figure 65 (SEQ ID NO:114), Figure 67 (SEQ ID NO:116), Figure 69 (SEQ ID NO:118), Figure 71 (SEQ ID NO:123), Figure 73 (SEQ ID NO:128), Figure 75 (SEQ ID NO:134), Figure 78 (SEQ ID NO:137), Figure 82 (SEQ ID NO:145), Figure 84 (SEQ ID NO:147), Figure 87 (SEQ ID NO:150), Figure 89 (SEQ ID NO:152), Figure 92 (SEQ ID NO:155), Figure 94 (SEQ ID NO:157), Figure 96 (SEQ ID NO:159), Figure 98 (SEQ ID NO:164), Figure 100 (SEQ ID NO:166), Figure 102 (SEQ ID NO:168), Figure 104 (SEQ ID NO:170), Figure 108 (SEQ ID NO:174), Figure 110 (SEQ ID NO:176), Figure 112 (SEQ ID NO:178), Figure 114 (SEQ ID NO:180), Figure 116 (SEQ ID NO:182), Figure 119 (SEQ ID NO:188), Figure 121 (SEQ ID NO:193), Figure 124 (SEQ ID NO:196), Figure 126 (SEQ ID NO:198), Figure 128 (SEQ ID NO:200), Figure 130 (SEQ ID NO:202), Figure 132 (SEQ ID NO:204), Figure 134 (SEQ ID NO:206), Figure 136 (SEQ ID NO:208), Figure 138 (SEQ ID NO:210), Figure 140 (SEQ ID NO:212), Figure 143 (SEQ ID NO:215), Figure 146 (SEQ ID NO:218), Figure 148 (SEQ ID NO:220), Figure 150 (SEQ ID NO:222), Figure 152 (SEQ ID NO:224), Figure 154 (SEQ ID NO:226), Figure 156 (SEQ ID NO:228), Figure 158 (SEQ ID NO:230), Figure 160 (SEQ ID NO:235), Figure 162 (SEQ ID NO:240), Figure 164 (SEQ ID NO:245), Figure 166 (SEQ ID NO:247), Figure 168 (SEQ ID

NO:249), Figure 170 (SEQ ID NO:252), Figure 173 (SEQ ID NO:255), Figure 175 (SEQ ID NO:257), Figure 177 (SEQ ID NO:259), Figure 179 (SEQ ID NO:261), Figure 181 (SEQ ID NO:263), Figure 183 (SEQ ID NO:265), Figure 185 (SEQ ID NO:267), Figure 187 (SEQ ID NO:269), Figure 189 (SEQ ID NO:271), Figure 191 (SEQ ID NO:273), Figure 193 (SEQ ID NO:275), Figure 195 (SEQ ID NO:277), Figure 197 (SEQ ID NO:280), Figure 199 (SEQ ID NO:282), Figure 201 (SEQ ID NO:284), Figure 203 (SEQ ID NO:286), Figure 205 (SEQ ID NO:288), Figure 207 (SEQ ID NO:290), Figure 209 (SEQ ID NO:292), Figure 211 (SEQ ID NO:294), Figure 213 (SEQ ID NO:296), Figure 215 (SEQ ID NO:298), Figure 217 (SEQ ID NO:300), Figure 219 (SEQ ID NO:302), Figure 225 (SEQ ID NO:308), Figure 227 (SEQ ID NO:313), Figure 229 (SEQ ID NO:318), Figure 232 (SEQ ID NO:325), Figure 234 (SEQ ID NO:333), Figure 237 (SEQ ID NO:339), Figure 239 (SEQ ID NO:344), Figure 241 (SEQ ID NO:346), Figure 243 (SEQ ID NO:348), Figure 245 (SEQ ID NO:350), Figure 247 (SEQ ID NO:352), Figure 249 (SEQ ID NO:354), Figure 251 (SEQ ID NO:356), Figure 253 (SEQ ID NO:358), Figure 255 (SEQ ID NO:360), Figure 257 (SEQ ID NO:362), Figure 259 (SEQ ID NO:364), Figure 261 (SEQ ID NO:366), Figure 263 (SEQ ID NO:368), Figure 265 (SEQ ID NO:370), Figure 267 (SEQ ID NO:372), Figure 269 (SEQ ID NO:374), Figure 271 (SEQ ID NO:376), Figure 273 (SEQ ID NO:378), Figure 275 (SEQ ID NO:380), Figure 277 (SEQ ID NO:386), Figure 279 (SEQ ID NO:388), Figure 281 (SEQ ID NO:393), Figure 283 (SEQ ID NO:398), Figure 285 (SEQ ID NO:400), Figure 287 (SEQ ID NO:402), Figure 289 (SEQ ID NO:407), Figure 291 (SEQ ID NO:409), Figure 293 (SEQ ID NO:411), Figure 295 (SEQ ID NO:413), Figure 297 (SEQ ID NO:415), Figure 299 (SEQ ID NO:417), Figure 301 (SEQ ID NO:419), Figure 303 (SEQ ID NO:421) and Figure 305 (SEQ ID NO:423).

- 20 23. An isolated extracellular domain of of PRO polypeptide.
24. An isolated PRO polypeptide lacking its associated signal peptide.
25. An isolated polypeptide having at least about 80% amino acid sequence identity to an
- 25 extracellular domain of of PRO polypeptide.
26. An isolated polypeptide having at least about 80% amino acid sequence identity to a PRO
- polypeptide lacking its associated signal peptide.

1/3-10

FIGURE 1

CGGACGCGTGGGTGCGAGGCGAAGGTGACCGGGGACCGAGCATTTTCAGATCTGCTCGGTAGA
CCTGGTGCACCACCACCATGTTGGCTGCAAGGCTGGTGTGTCTCCGGACACTACCTTCTAGG
GTTTTCCACCCAGCTTTCACCAAGGCCTCCCCTGTTGTGAAGAATTCCATCACGAAGAATCA
ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAAACAAGAATTGGGATCCGGCGTGGGA
GAACTGGCCAAGAACTCAAAGAGGCAGCATTGGAACCATCGATGGAAAAAATATTTAAAATT
GATCAGATGGGAAGATGGTTTGTGCTGGAGGGGCTGCTGTTGGTCTTGGAGCATTGTGCTA
CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT
ATGTCAAGGATAGAATTCATTCCACCTATATGTACTTAGCAGGGAGTATTGGTTTAAACAGCT
TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACTTCATGATGAGAGGCTCTTG
GGTGACAATTGGTGTGACCTTTGCAGCCATGGTTGGAGCTGGAATGCTGGTACGATCAATAC
CATATGACCAGAGCCCAGGCCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT
GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCTCTTCTCATCAGAGCTGCATGGTACAC
AGCTGGCATTGTGGGAGGCCTCTCCACTGTGGCCATGTGTGCGCCCAGTGAAAAGTTTCTGA
ACATGGGTGCACCCCTGGGAGTGGGCCTGGGTCTCGTCTTTGTGTCTTCATTGGGATCTATG
TTTCTTCCACCTACCACCGTGGCTGGTGCCACTCTTTACTCAGTGGCAATGTACGGTGGATT
AGTTCTTTTTCAGCATGTTCTCTTCTGTATGATACCCAGAAAGTAATCAAGCGTGCAGAAGTAT
CACCAATGTATGGAGTTCAAAAATATGATCCCATTAACTCGATGCTGAGTATCTACATGGAT
ACATTAAATATATTTATGCGAGTTGCAACTATGCTGGCAACTGGAGGCAACAGAAAGAAATG
AAGTGACTCAGCTTCTGGCTTCTCTGCTACATCAAATATCTTGTTTAAATGGGGCAGATATGC
ATTAAATAGTTTGTACAAGCAGCTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA
TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTTCTGGAGAATAAATGCAGT
AATCCTCTCCCAAATAAGCACACACATTTTCAATTCTCATGTTTGAGTGATTTTAAATGTT
TTGGTGAATGTGAAAATAAGTTTGTGTCATGAGAATGTAAGTCTTTTTTCTACTTTAAAA
TTTAGTAGGTTCACTGAGTAACTAAAATTTAGCAAACCTGTGTTTGCATATTTTTTTGGAGT
GCAGAATATTGTAATTAATGTCATAAGTGATTGGAGCTTTGGTAAAGGGACCAGAGAGAAG
GAGTCACCTGCAGTCTTTTGTTTTTTTTAAATACTTAGAACTTAGCACTTGTTATTGATTA
GTGAGGAGCCAGTAAGAAACATCTGGGTATTTGGAAACAAGTGGTCATTGTTACATTCAATT
GCTGAACTTAACAAAACCTGTTTCATCCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG
CTTCTCAGTGCTCTCTTTCCAATATAGATGTGGTCATGTTTGACTTGTACAGAATGTTAATC
ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTTACTTTTGAATGTTACAAAAGGAA
ATAACTTTAAACTATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTCCAG
AATACAAACAGTATACTCATG

FIGURE 2

MLAARLVCLRTLPSRVFHPAFTKASPVVKNSITKNQWLLTPSREYATKTRIGIRRGRTGQEL
KEAALEPSMEKIFKIDQMGRWVAVAGGAAVGLGALCYGLGLSNEIGAIEKAVIWPQYVKDRI
HSTYMYLAGSIGLTALSAIAISRTPVLMNFMGRGSWVTIGVTFAAMVGAGMLVRSIPYDQSP
GPKHLAWLLHSGVMGAVVAPLTILGGPLLIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL
GVGLGLVFVSSLGSMFLPPTTVAGATLYSVAMYGGLVLFMFLLYDTQKVIKRAEVSPMYGV
QKYDPINSMLSIYMDTLNIFMRVATMLATGGNRKK

FIGURE 3

GAAGGCTGCCTCGCTGGTCCGAATTCGGTGGCGCCACGTCCGCCCCTCTCCGCCTTCTGCAT
CGCGGCTTCGGCGGCTTCCACCTAGACACCTAACAGTCCGCGGAGCCGGCCGCGTCTGAGGG
GGTCCGGCACGGGGAGTCCGGCGGTCTTGTGCATCTTGGCTACCTGTGGGTCTGAAGATGTCGG
ACATCGGAGACTGGTTTCAGGAGCATCCCGGCGATCACGCGCTATTGGTTCGCCCCCACCGTC
GCCGTGCCCTTGGTCCGCAAACCTCGGCCTCATCAGCCCGGCCTACCTCTTCCTCTGGCCCGA
AGCCTTCCTTTATCGCTTTCAGATTTGGAGGCCAATCACTGCCACCTTTTATTTCCCTGTGG
GTCCAGGAACCTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGA
CTTGAAACAGGAGCTTTTGATGGGAGGCCAGCAGACTATTTATTCATGCTCCTCTTTAACTG
GATTTGCATCTTGATTACTGGCTTAGCAATGGATATGCAGTTGCTGATGATTCCTCTGATCA
TGTCAGTACTTTATGTCTGGGCCAGCTGAACAGAGACATGATTGTATCATTTTGGTTTGGA
ACACGATTTAAGGCCTGCTATTTACCCTGGGTTATCCTTGGAATCAACTATATCATCGGAGG
CTCGGTAATCAATGAGCTTATTGGAAATCTGGTTGGACATCTTTATTTTTTCTTAATGTTCA
GATACCCAATGGACTTGGGAGGAAGAAATTTCTATCCACACCTCAGTTTTTGTACCGCTGG
CTGCCCAGTAGGAGAGGAGGAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGC
TGCTGATCAGAATGGCGGAGGCGGGAGACACAACCTGGGGCCAGGGCTTTCGACTTGGAGACC
AGTGAAGGGGCGGCCTCGGGCAGCCGCTCCTCTCAAGCCACATTTCTCCAGTGCTGGGTG
CACTTAACAACCTGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTCTTTC
AGTACGAGACAAAGTTTCTTAAATCCCGAAGAAAAATATAAGTGTTCACAAGTTTCACGAT
TCTCATTCAAGTCTTACTGCTGTGAAGAACAAATACCAACTGTGCAAATTGCAAACTGAC
TACATTTTTTGGTGTCTTCTCTTCTCCCCTTCCGTCTGAATAATGGGTTTTAGCGGGTCTT
AATCTGCTGGCATTGAGCTGGGGCTGGGTCAACCAACCTTCCCAAAGGACCTTATCTCTT
TCTTGACACATGCCTCTCTCCACTTTTCCCAACCCCCACATTTGCAACTAGAAAAAGTTG
CCCATAAAATTGCTCTGCCCTTGACAGGTTCTGTTATTTATTGACTTTTGCCAAGGCTGGTC
ACAACAATCATATTACGTTATTTTCCCCTTTTGGTGGCAGAACTGTTACCAATAGGGGGAG
AAGACAGCCACGGATGAAGCGTTTCTCAGCTTTTGGAAATGCTTCGACTGACATCCGTTGTT
AACCGTTTGCCACTCTTCAGATATTTTTTATAAAAAAGTACCACTGAGTTCATGAGGCCCA
CAGATTGGTTATTAATGAGATACGAGGGTTGGTGTGGGTGTTTGTTCCTGAGCTAAGTGA
TCAAGACTGTAGTGGAGTTGCAGCTAACATGGGTTAGGTTTAAACCATGGGGGATGCACCCC
TTTGCGTTTTCATATGTAGCCCTACTGGCTTTGTGTAGCTGGAGTAGTTGGGTTGCTTTGTGT
TAGGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGTCCTGGGCATTG
ATTCCCATTTCATCTCATTCTGGATATGTGTTCAATTGAGTAAAGGAGGAGAGACCCTCATA
CGCTATTTAAATGTCACTTTTTTGCCTATCCCCCGTTTTTGGTTCATGTTTCAATTAATTGT
GAGGAAGGCGCAGCTCCTCTCTGCACGTAGATCATTTTTTAAAGCTAATGTAAGCACATCTA
AGCGAATAACATGATTTAAGGTTGAAATGGCTTTAGAATCATTTGGGTTTGAGGGTGTGTTA
TTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTT
TCGTAGGTGGGCTTTTCCTATCAGAGCTTGGCTCATAACCAAATAAAGTTTTTTGAAGGCCA
TGGCTTTTTCACACAGTTATTTTATTTTATGACGTTATCTGAAAGCAGACTGTTAGGAGCAGT
ATTGAGTGGCTGTCACTTTTGAGGCAACTAAAAAGGCTTCAAACGTTTTGATCAGTTTCTT
TTCAGGAAACATTGTGCTCTAACAGTATGACTATTTCTTTCCCCCACTCTTAAACAGTGTGAT
GTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACCTTCTCATTTTGAATAGAGTTTGTGTG
TACTTCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAACTGTCA
TGTTTTGTTGTTTCATCTGTGGCCACAATAAAGTTTACTTGTAATAATTTAGAGGCCATTACT
CCAATTATGTTGCACGTACACTCATTGTACAGGCGTGGAGACTCATTGTATGTATAAGAATA
TTTCTGACAGTGAGTGACCCGGAGTCTCTGGTGTACCCTCTTACCAGTCAGCTGCCTGCGAG
CAGTCATTTTTTCTTAAAGGTTTACAAGTATTTAGAATTTTTCAGTTTCAGGGCAAAATGTTT
ATGAAGTTATTCCTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTGTCTGGATT
ATGTTTCTGGAATAATTTTACCAAAACAAGCTATTTGAGTTTGTACTTGACAAGGCAAAACA
TGACAGTGGATTCTCTTTACAAATGGAAAAAAAATCCTTATTTTGTATAAAGGACTTCCC
TTTTTGTAACATAATCCTTTTTTATTGGTAAAAATTGTAAATTAAATGTGCAACTTG

FIGURE 4

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGLISPAYLFLWPEAFLYRFQIWRPITATFYF
PVGPGTGFLYLVNLYFLYQYSTRLETGAFDGRPADYLEMLLFNWICIVITGLAMDMQLLMIP
LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIGGSVINELIGNLVGHLYFFL
MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRLGDQ

FIGURE 5

GGGGCCGCGGTCTAGGGCGGCTACGTGTGTTGCCATAGCGACCATTTTGCATTAACTGGTTG
GTAGCTTCTATCCTGGGGGCTGAGCGACTGCGGGCCAGCTCTTCCCCTACTCCCTCTCGGCT
CCTTGTGGCCCAAAGGCCTAACCAGGGGTCCGGCGGTCTGGCCTAGGGATCTTCCCCGTTGCC
CCTTTGGGGCGGGATGGCTGCGGAAGAAGAAGACGAGGTGGAGTGGGTAGTGGAGAGCATCG
CGGGGTTTCTGCGAGGCCCAGACTGGTCCATCCCCATCTTGGACTTTGTGGAACAGAAATGT
GAAGTTAACTGCAAAGGAGGGCATGTGATAACTCCAGGAAGCCCAGAGCCGGTGATTTTGGT
GGCCTGTGTTCCCCTTGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGATTC
ATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAATT
AATGAAGATCAATTTCAAGAAGCATGCACCTTCTCCTCTTGCAAAGACCCATACATCACAGGC
CATTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCCAGA
AAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAATGGTGTATTACCT
GACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAATCCT
GAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATGACCAGGAAGAAGAAAGGAAGAGGAAAA
AACAGTTATCAGAGGCTAAAACAGAAGAGCCACAGTGCATTCCAGTGAAGCTGCAATAATG
AATAATTCCCAAGGGGATGGTGAACATTTTGCACACCCACCCTCAGAAGTTAAAATGCATTT
TGCTAATCAGTCAATAGAACCTTTGGGAAGAAAAGTGGAAAGGTCTGAAACTTCCTCCCTCC
CACAAAAGGCCTGAAGATTCTTGCTTAGAGCATGCGAGCATTGAAGGACCAATAGCAAAC
TTATCAGTACTTGGAACAGAAGAACTTCGGCAACGAGAACACTATCTCAAGCAGAAGAGAGA
TAAGTTGATGTCCATGAGAAAGGATATGAGGACTAAACAGATACAAAATATGGAGCAGAAAG
GAAAACCCACTGGGGAGGTAGAGGAAATGACAGAGAAACCAGAAATGACAGCAGAGGAGAAG
CAAACATTACTAAAGAGGAGATTGCTTGCAGAGAACTCAAAGAAGAAGTTATTAATAAGTA
ATAATTAAGAACAATTTAACAAAATGGAAGTTCAAATTGTCTTAAAAATAAATTATTTAGTC
CTTACACTG

6/310

FIGURE 6

MAAEEDEVEWVESIAGFLRGPDWSIPILDFVEQKCEVNCKGGHVITPGSPEPVILVACVP
LVFDDEEESKLTYTEIHQEYKELVEKLLEGYLKEIGINEDQFQEAactsPLAKTHTSQAILQP
VLAAEDFTIFKAMMVQKNIEMQLQAIRIIQERNGLVLPDCLTDGSDVVSdleHEEMKILREVL
RKSKEEYDQEEERKRKKQLSEAKTEEPTVHSSEAAIMNNSQGDGEHFAHPPSEVKMHFANQS
IEPLGRKVERSETSSLPQKGLKIPGLEHASIEGPiANLSVLGTEELRQREHYLKQKRDKLMS
MRKDMRTKQIQNMEQKGKPTGEVEEMTEKPEMTAEKQTLLKRLLAEKLKEEVINK

FIGURE 7

GGGCACAGCACATGTGAAGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGAT
TCATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAA
TTAATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAG
GCCATTTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCC
AGAAAAACATTGAAATGCAGCTGCAAGCCATTTCGAATAATTCAAGAGAGAAATGGTGTATTA
CCTGACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAAT
CCTGAGGGAAGTTCTTAGAAAATCAAAGAGGAATATGACCAGGAA

FIGURE 8

GC GTG G T T T T T G T T C T G C A A T A G G C G G C T T A G A G G G A G G G G C T T T T T C G C C T A T A C C T A C T G
T A G C T T C T C C A C G T A T G G A C C C T A A A G G C T A C T G C T A C T A C G G G G C T A G A C A G T T A C T G
T C T C A G C T C T A G G A T G T G C G T T C T T C C A C T A G A A G C T C T T C T G A G G G A G G T A A T T A A A A A C
A G T G G A A T G G A A A A A C A G T G C T G T A G T C A T C C T G T A A T A T G C T C C T T G T C A A C A A T G T A T A C
A T T C C T G C T A G G T G C C A T A T T C A T T G C T T T A A G C T C A A G T C G C A T C T T A C T A G T G A A G T A T T
C T G C C A A T G A A G A A A C A A G T A T G A T T A T C T T C C A A C T A C T G T G A A T G T G T G C T C A G A A C T G
G T G A A G C T A G T T T T C T G T G T G C T T G T G T C A T T C T G T G T T A T A A G A A A G A T C A T C A A A G T A G
A A A T T G A A A T A T G C T T C C T G G A A G G A A T T C T C T G A T T T C A T G A A G T G G T C C A T T C C T G C C T
T T C T T T A T T T C C T G G A T A A C T T G A T T G T C T T C T A T G T C C T G T C C T A T C T T C A A C C A G C C A T G
G C T G T T A T C T T C T C A A A T T T T A G C A T T A T A A C A A C A G C T C T T C T A T T C A G G A T A G T G C T G A A
G A G G C G T C T A A A C T G G A T C C A G T G G G C T T C C C T C C T G A C T T T A T T T T G T C T A T T G T G G C C T
T G A C T G C C G G G A C T A A A A C T T T A C A G C A C A A C T T G G C A G G A C G T G G A T T T C A T C A C G A T G C C
T T T T T C A G C C C T T C C A A T T C C T G C C T T C T T T T C A G A A G T G A G T G T C C C A G A A A A G A C A A T T G
T A C A G C A A A G G A A T G G A C T T T T C C T G A A G C T A A A T G G A A C A C C A C A G C C A G A G T T T T C A G T C
A C A T C C G T C T T G G C A T G G G C C A T G T T C T T A T T A T A G T C C A G T G T T T T A T T T C T T C A A T G G C T
A A T A T C T A T A A T G A A A G A T A C T G A A G G A G G G G A A C C A G C T C A C T G A A A G C A T C T T C A T A C A
G A A C A G C A A A C T C T A T T T C T T T G G C A T T C T G T T T A A T G G G C T G A C T C T G G G C C T T C A G A G G A
G T A A C C G T G A T C A G A T T A A G A A C T G T G G A T T T T T T A T G G C C A C A G T G C A T T T T C A G T A G C C
C T T A T T T T T G T A A C T G C A T T C C A G G G C C T T T C A G T G G C T T T C A T T C T G A A G T T C C T G G A T A A
C A T G T T C C A T G T C T T G A T G G C C C A G G T T A C C A C T G T C A T T A T C A C A A C A G T G T C T G T C C T G G
T C T T T G A C T T C A G G C C C T C C C T G G A A T T T T C T T G G A A G C C C A T C A G T C C T T C T C T A T A
T T T A T T T A T A A T G C C A G C A A G C C T C A A G T T C C G G A A T A C G C A C C T A G G C A A G A A A G G A T C C G
A G A T C T A A G T G G C A A T C T T T G G G A G C G T T C C A G T G G G G A T G G A G A A G A A C T A G A A A G A C T T A
C C A A A C C C A A G A G T G A T G A G T C A G A T G A A G A T A C T T T C T A A A C T G G T A C C C A C A T A G T T T G C A
G C T C T C T T G A A C C T T A T T T T C A C A T T T T C A G T G T T G T A A T A T T T A T C T T T T C A C T T T G A T A
A A C C A G A A A T G T T T C T A A A T C C T A A T A T T C T T T G C A T A T A T C T A G C T A C T C C C T A A A T G G T T
C C A T C C A A G G C T T A G A G T A C C C A A A G G C T A A G A A A T T C T A A A G A A C T G A T A C A G G A G T A A C A
A T A T G A A G A A T T C A T T A A T A T C T C A G T A C T T G A T A A A T C A G A A A G T T A T A T G T G C A G A T T A T
T T T C C T T G G C C T T C A A G C T T C C A A A A A C T T G T A A T A A T C A T G T T A G C T A T A G C T T G T A T A T
A C A C A T A G A G A T C A A T T T G C C A A A T A T T C A C A A T C A T G T A G T T C T A G T T T A C A T G C C A A A G T
C T T C C C T T T T T A A C A T T A T A A A A G C T A G G T T G T C T C T T G A A T T T T G A G G C C C T A G A G A T A G T
C A T T T T G C A A G T A A A G A G C A A C G G G A C C C T T T C T A A A A C G T T G G T T G A A G G A C C T A A A T A C
C T G G C C A T A C C A T A G A T T T T G G G A T G A T G T A G T C T G T G C T A A A T A T T T T G C T G A A G A A G C A G T
T T C T C A G A C A C A A C A T C T C A G A A T T T T A A T T T T T A G A A A T T C A T G G G A A A T T G G A T T T T T G T
A A T A A T C T T T T G A T G T T T T A A C A T T G G T T C C C T A G T C A C C A T A G T T A C C A C T T G T A T T T T A
A G T C A T T T A A A C A A G C C A C G G T G G G G C T T T T T T C T C C T C A G T T T G A G G A G A A A A T C T T G A T
G T C A T T A C T C C T G A A T T A T T A C A T T T T G G A G A A T A A G A G G G C A T T T T A T T T T A T T A G T T A C T
A A T T C A A G C T G T G A C T A T T G T A T A T C T T T C C A A G A G T T G A A A T G C T G G C T T C A G A A T C A T A C
C A G A T T G T C A G T G A A G C T G A T G C C T A G G A A C T T T T A A A G G G A T C C T T T C A A A A G G A T C A C T T
A G C A A A C A C A T G T T G A C T T T T A A C T G A T G T A T G A A T A T T A A T A C T C T A A A A A T A G A A A G A C C
A G T A A T A T A T A A G T C A C T T T A C A G T G C T A C T T C A C A C T T A A A A G T G C A T G G T A T T T T T C A T G
G T A T T T T G C A T G C A G C C A G T T A A C T C T C G T A G A T A G A A A G T C A G G T G A T A G A T G A T A T T A A
A A A T T A G C A A A C A A A A G T G A C T T G C T C A G G G T C A T G C A G C T G G G T G A T G A T A G A A G A G T G G G
C T T T A A C T G G C A G G C C T G T A T G T T T A C A G A C T A C C A T A C T G T A A A T A T G A G C T T T A T G G T G T
C A T T C T C A G A A C T T A T A C A T T T T C T G C T C C T T T C C T A A G T T T C A T G C A G A T G A A T A T A
A G G T A A T A T A C T A T T A T A T A A T T C A T T T G T G A T A T C C A C A A T A A T A T G A C T G G C A A G A A T T G
G T G G A A A T T T G T A A T T A A A A T A A T T A T T A A A C C T

FIGURE 9

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSSELVK
LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQPAMAV
IFS NFSIITTALLFRIVLKRRLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFF
SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFSHIRLGMGHVLIIVQCFISSMANI
YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFYGHSAFSVALI
FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI
YNASKPQVPEYAPRQERIRDLSGNLWERSSSGDGEELERLTKPKSDESDTF

FIGURE 10

CGTGCCTGCGCAATGGGTGTCGGGTCCGCTTTTTCCCAATCCGGACGTAATCGTGGTTTTTG
TTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTCGCCTATACCTACTGTAGCTTCTCCAC
GTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTGTCTCAGCTCTAG
GATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTAAAAAACAGTGGAATGGAA
AAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCACAATGTATACATTCCTGCTAGG
TGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATTCTGCCAATGAAG
AAAACAAGTATGATTATCTTCCAACACTACTGTGAATGTGTGCTCAGAACTGGTGAAGCTAGTT
TTCTGTGTGCTTGTGTCATTCTGTGTTATAAAGAAAGATCATCAAAGTAGAAATTTGAAATA
TGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCCTGCCTTTCTTTATTTCC
TGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATGGCTGTTATCTTC
TCAAATTTTAGCATTATAACAACAGCTCTTCTATTCAGGATAGTGCTGAAGAGGCGTCTAAA
CTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCTTGACTGCCGGGA
CTAAAACTTTA

FIGURE 11

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGGCCGGCTTGGCTAGCGCGCGGCGGCC
GTGGCTAAGGCTGCTACGAAGCGAGCTTGGGAGGAGCAGCGGCCTGCGGGGACAGAGGAGCAT
CCCGTCTACCAGGTCCCAAGCGGCGTGGCCCGCGGGTCATGGCCAAAGGAGAAGGCGCCGAG
AGCGGCTCCGCGGCGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCGGCCCA
GGTGAAGAAAGAACCAGAAAAAGAAACAACAGTTGTCTGTTTGCAACAAGCTTTGCTATG
CACTTGGGGGAGCCCCCTACCAGGTGACGGGCTGTGCCCTGGGTTTCTTCCTTCAGATCTAC
CTATTGGATGTGGCTCAGGTGGGCCCTTTCTCTGCCTCCATCATCCTGTTTGTGGGCCGAGC
CTGGGATGCCATCACAGACCCCTGGTGGGCCTCTGCATCAGCAAATCCCCCTGGACCTGCC
TGGGTGCGCTTATGCCCTGGATCATCTTCTCCACGCCCCCTGGCCGTCATTGCCTACTTCCTC
ATCTGGTTTCGTGCCCAGCTTCCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCT
CTTTGAAACAATGGTCACGTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTTCATCAGCA
ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC
AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCTTGTTTCCAGG
ACTTCAATAGCTCTACAGTAGCTTCACAAAGTGCCAACCATACACATGGCACCCTTCACAC
AGGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGTCAATTGTCTGTATCTATATAATCTG
TGCTGTATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCCTATGAAGCCAGCAGTCTG
AGCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACTCAAACCTT
ATTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACCTTTGTCTTGTT
TTGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCT
CGGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCT
GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAA
CCTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC
TACCCTGGTCCATGCTGCCTGATGTGATGACGACTTCCATCTGAAGCAGCCCCACTTCCAT
GGAACCGAGCCCATCTTCTTCTCCTTCTATGTCTTCTTACCAAGTTTGCCTCTGGAGTGTG
ACTGGGCATTTCTACCCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC
CGGAACGTGTCAAGTTTACACTGAACATGCTCGTGACCATGGCTCCCATAGTTCTCATCCTG
CTGGGCCTGCTGCTCTTCAAAATGTACCCCATGATGAGGAGAGGCGGCGGCAGAATAAGAA
GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTGCTCAGAAACAGACTCCACAG
AGCTGGCTAGCATCCTCTAGGGCCCGCCACGTTGCCCGAAGCCACCATGCAGAAGGCCACAG
AAGGGATCAGGACCTGTCTGCCGGCTTGCTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA
CTGAAGACTCAAGGAGGTGGCCCAGGACACTTGCTGTGCTCACTGTGGGGCCGGCTGCTCTG
TGGCCTCCTGCCTCCCCTCTGCCTGCCTGTGGGGCCAAGCCCTGGGGCTGCCACTGTGAATA
TGCCAAGGACTGATCGGGCCTAGCCCGGAACACTAATGTAGAAACCTTTTTTTTACAGAGCC
TAATTAATAACTTAATGACTGTGTACATAGCAATGTGTGTGTATGTATATGTCTGTGAGCTA
TTAATGTTATTAATTTTCATAAAAGCTGGAAAGC

FIGURE 12

MWLRWALSLPPSSCLWAEPMPSQTPWWASASANPPGPAWVALCPGSSSPRPWPSLPTSSSG
SCPTSHTARPIGTCSIASLKQWSRVSMFPTRLSPCSSATEQTERDSATAYRMTVEVLGTVL
GTAIQGQIVGQADTPCFQDFNSSTVASQSANTHGTTSHRETQKAYLLAAGVIVCIYIICAV
ILILGVREQREPYEAQQSEPIAYFRGLRLVMSHGPYIKLITGFLFTSLAFMLVEGNFVLFCT
YTLGFRNEFQNLLLAIMLSATLTIPWQWFLTRFGKKTAVYVGISSAVPFLILVALMESNLI
ITYAVAVAAGISVAAAFLLPWSMLPDVIDDFHLKQPHFHGTEPIFFSFYVFFTKFASGVSLG
ISTLSLDFAGYQTRGCSQPERVKFTLNMLVTMAPIVLILLGLLLFKMYPIDEERRRQNKAL
QALRDEASSSGCSETDSTELASIL

FIGURE 13

GGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT
GCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTGA
GCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACCTTA
TTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACCTTGTCTTGTTT
TGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCTC
GGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCTG
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAAC
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT
ACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG
GAACCGAGCCCCAT

FIGURE 14

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGT
ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGT
GGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT
TTCATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT
CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGG
CAGTTTTATGCATTGCTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA
GAGAACGTTATCATCAAATTAACAAGGCTGGCCTTGTAAGTCTGGAATACTGAGTTGTTTAGG
ACTTTCTATTGTGGCAAACCTTCCAGAAAACAACCCCTTTTGCTGCACATGTAAGTGGAGCTG
TGCTTACCTTTGGTATGGGCTCATTATATATGTTTGTTCAGACCATCCTTTCCTACCAAATG
CAGCCCCAAAATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGG
AGTAAGTGCACTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTGGGACTG
ATTTAGAACAGAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACT
ACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTTGGTTTTTTCCTGACTTACATTCGTGA
TTTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCCTCTATGACACTG
CACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCAGAGATATTTGATGAAAGGAT
AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTACAGAAAGTTGCTTA
TTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA
ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA
TTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACTATG

FIGURE 15

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNI
AAVLCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSG
AVLTFGMGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFG
TDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYD
TAPCPINNERTRLLSRDI

FIGURE 16

CGGACGCTTGGGCNGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGTGCCTGATGCCGAGT
TCCGTCTCTCGGGTCTTTTCCTGGTCCCAGGCAAAGCGGAGCGGAGATCCTCAAACGGCCTA
GTGCTTCGCGCTTCCGGAGAAAATCAGCGGTCTAATTAATTCCTCTGGTTTGTTGAAGCAGT
TACCAAGAATCTTCAACCCTTTCCCACAAAAGCTAATTGAGTACACGTTCCCTGTTGAGTACA
CGTTCCTGTTGATTTACAAAAGGTGCAGGTATGAGCAGGTCTGAAGACTAACATTTTGTGAA
GTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCT
TCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTTTCATACATTACTGCAGTAACACT
CCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGGTACAGTANC

FIGURE 17

CCCACGCGTCCGCCCCGCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCCGCGGG
CCGGGGTGCGGAGCCGACCATGCGCCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCAC
CTTCGCCTTGTAATTGCTGTGACGCGACTGCCCCGCGGGCGGAGACTGGGCTCCACCGAGG
AGGCTGGAGGCAGGTGCTGTGGTTCCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG
GTCCTTCGAGAGTACCGGAAGGAGCACCAGGCCTACGTGTTTCTGCTCTTCTGCGGCGCCTA
CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCTGAATGTTTTAGCTGGTGCCT
TGTTTGGGCCATGGCTGGGGCTTCTGCTGTGCTGTGTGTTGACCTCGGTGGGTGCCACATGC
TGCTACCTGCTCTCCAGTATTTTTGGCAAACAGTTGGTGGTGTCTACTTTCTGATAAAGT
GGCCCTGCTGCAGAGAAAGGTGGAGGAGAACAGAAACAGCTTGTTTTTTTCTTATTGTTTT
TGAGACTTTTTCCCATGACACCAAACCTGGTTCTTGAACCTCTCGGCCCCAATTCTGAACATT
CCCATCGTGCAGTTCTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTTTCATCTGTGT
GCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTG
TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAATTCCTGGAACCCTCATTAAAAAATTT
AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGA
CACATGATCTGGATTTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA
TGTGGTCTCTAAAGCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG
CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT
TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGT
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATTCT
ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAT
ACAAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGC
AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT
CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

FIGURE 18

MRPLLGLLLVFAGCTFALYLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAELRELSEVLREYR
KEHQAYVFLLFCGAYLYKQGFAIPGSSFLNVLAGALFGPWLGLLLCCVLTSVGATCCYLLSS
IFGKQLVVSYFPDKVALLQRKVEENRNSLFFFLFLRLFPMTPNWFLNLSAPILNIPIVQFF
FSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLLAIAMVALIPGTLIKKFSQKHLQ
LNETSTANHIHSRKDT

FIGURE 19

CCGAGGCGGGAGGAGCCCCGAGGGGGCGCGAGCCCCGCATGAATCATTGTAGTCAATCATTTT
CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAA
TAGGAAAATAACTTGGGATTTTATATTGGAAGACATGGATCTTGCTGCCAACGAGATCAGCA
TTTATGACAACTTTCAGAGACTGTTGATTTGGTGAGACAGACCGGCCATCAGTGTGGCATG
TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC
CCCCCGCAGTATCCTCTCCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT
TGCTCACTGCCTACTTTGTGATTCAACCTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTCT
GGAGCTCACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCCTTGCCCATTGCCAA
GAAGTACATGTCAGAAAATAAGGGAGTTCCTCTGCATGGGGGTGATGAAGACAGACCCTTTC
CAGACTTTGACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCTCTGCC
AACTGCACTGGCTGTGCCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA
ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTC
AGCATTTTTTGTGCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTTCGCCAAG
TGGTGGCGCTGCTTTCCTGAGCGGTGGTTCCCATTTCTTATCCATGGAGGAGACCTCTGAA
CAGATCACAAATGTTACGTGAGCTTTTTCTGTTTTCACTCACCTGCCATTTCCAAAAGATG
CCTCTTTAAACAAGTGCTCCTTTCTTCACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAG
ATGCCTGACCTATTTATCATTTGGCAGCGGTGAGGCCATGTTGCAGCTCATCCCTCCCTTCCA
GTGCCGAAGACATTGTCACTGTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTCTG
ACACCACCCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT
GGAACCGCTTTCTCAGAACTGTAGGAAATAGAACTGTGCACAGGAACAGCTTCCAGAGCCGA
AAACCAGGTTGAAAGGGGAAAAATAAAAAACAAAACGATGAAACTGCAAAAA

FIGURE 20

MDLAANEISIIYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPPQYPLLIVVY
KVLATLGLILLTAYFVIQPFSPPLAPEPVLSGAHTWRSLIHHIRLMSLPPIAKKYMSENKGVPL
HGGDEDRPFPDFDPWWTNDCEQNESEPI PANCTGCAQKHLKVMLLEDAPRKFERLHPLVIKT
GKPLLEEEIQHFLCQYPEATEGFSEGFFAKWWRCFPERWFPPYPWRRPLNRSQMLRELFV
FTHLPFPKDASLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQCRRHCSVAMP
IEPGDIGYVDTTHWKVYVIARGVQPLVICDGTAFSEL

FIGURE 21

CCACGGTGTCCGTTCTTCGCCCCGGCGGCAGCTGTCCCCGAGGCGGGAGGAGCCCCGAGGGGCG
CGAGCCCCGCATGAATCATTGTAGTCAATCATTTTCCAGTTCTCAGCCGTTCAAGTTGTGATC
AAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAATAGGAAAATAAAGTTGGGATTTTATATT
GGAAGACATGGATCTTGCTGCCAACGAGATCAGCATTTATGACAAAGTTTCAGAGACTGTTG
ATTTGGTGAGACAGACCGGCCATCAGTGTGGCATGTCAGAGAAGGCAATTGAAAAATTTATC
AGACAGCTGCTGGAAAAGAATGAACCTCAGAGACCCCCCGCAGTATCCTCTCCTTATAGT
TGTGTATAAGGTTCTCGCAACCTTGGGATTAATCTTGCTCACTGCCTACTTTGTGATTCAAC
CTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTGTGGAGCTCAC

FIGURE 22

CCCCACGCGTCCGCCCCACGCGTCCGGCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGG
CAGGGTCCCCACTTGCAGCTGCAGCAGCTGCAGCAGCTGCAGAGCGCTGCTCCTGGCTGGTG
CCACTGGTGCACGCTGCTAGACCGTGCCTATGAGCCGCTGGGGCTGCAGTGGGGACTGCC
CTCCCTGCCACCCACCAATGGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTGTGCCA
CACCCGAATGGCGCCACTTCATCGACAAACAGGTACAGCCAACCATGTCCAGTTTCGAAATG
GACACGTATGCTAAGAGCCACGACCTTATGTAGGTTTTCTGGAATGCCTGCTATGACATGCT
TATGAGCAGTGGGCAGCGGCGCCAGTGGGAGCGCGCCAGAGTCGTCGGGCCTTCCAGGAGC
TGGTGCTGGAACCTGCGCAGAGGCGGGCGCGCTGGAGGGGCTACGCTACACGGCAGTGCTG
AAGCAGCAGGCAACGCAGCACTCCATGGCCCTGCTGCACTGGGGGGCGCTGTGGCGCCAGCT
CGCCAGCCCATGTGGGGCCTGGGCGCTGAGGGACACTCCCATCCCCGCTGGAAACTGTCCA
GCGCCGAGACATATTACGCATGCGTCTGAAGCTGGTGCCCAACCATCACTTCGACCCTCAC
CTGGAAGCCAGCGCTCTCCGAGACAATCTGGGTGAGGTTCCCCTGACACCCACCGAGGAGGC
CTCACTGCCTCTGGCAGTGACCAAAGAGGGCAAAGTGAGCACCCCCACCCGAGTTGCTGCAGG
AGGACCAGCTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCCGATGGAGGCAGCAGAACTG
GATGAGCAGCGTGAGAAGCTGGTGCTGTGGCCGAGTGCCAGCTGGTGACGGTAGTGCCCGT
GGTCCCAGGGCTGCTGGAGGTACCCACACAGAATGTATACTTCTACGATGGCAGCACTGAGC
GCGTGGAACCGAGGAGGGCATCGGCTATGATTTCCGGCGCCCACTGGCCAGCTGCGTGAG
GTCCACCTGCGGCGTTTCAACCTGCGCCGTTGAGCACTTGAGCTCTTCTTTATCGATCAGGC
CACTACTTCTCAACTTCCCATGCAAGGTGGGCACGACCCAGTCTCATCTCTCAGGAGCA
CTCCGAGACCCAGCCTGGCCCCATCCCCACCCATACCCAGGTACGGAACCGAGGTGTAAGT
TGGCTCCTGCGCTACGGCCCCCTCTCAAGGCTACCTAAGCAGCCGCTCCCCCAGGAGAT
GCTGCGTGCTCAGGCCTTACCCAGAAATGGGTACAGCGTGAGATATCCAACCTTCGAGTACT
TGATGCAACTCAACACCATTGCGGGGCGGACCTACAATGACCTGTCTCAGTACCCTGTGTTC
CCCTGGGTCTGACGAGTACGTGTCCCCAACCCCTGGACCTCAGCAACCCAGCCGTCTTCCG
GGACCTGTCTAAGCCCATCGGTGTGGTGAACCCCAAGCATGCCAGCTCGTGAGGGAGAAGT
ATGAAAGCTTTGAGGACCCAGCAGGGACCATTGACAAGTTCCACTATGGCACCCACTACTCC
AATGCAGCAGGCGTGATGCACTACCTCATCCGCGTGGAGCCCTTACCTCCCTGCACGTCCA
GCTGCAAAGTGGCCGCTTTGACTGCTCCGACCGGCAGTTCCACTCGGTGGCGGCAGCCTGGC
AGGCACGCCTGGAGAGCCCTGCCGATGTGAAGGAGCTCATCCCGGAATTCTTCTACTTTCT
GACTTCTGGAGAACCAGAACGGTTTTGACCTGGGCTGTCTCCAGCTGACCAACGAGAAGGT
AGGCGATGTGGTGCTACCCCCGTGGGCCAGCTCTCTGAGGACTTCATCCAGCAGCACC GCC
AGGCTCTGAGAGTCGAGTATGTGTCTGCACACCTACAGAGTGGATCGACCTCATCTTTGGC
TACAAGCAGCGGGGGCCAGCCGCGGAGGAGCCCTCAATGTCTTCTATTACTGCACCTATGA
GGGGGCTGTAGACCTGGACCATGTGACAGATGAGCGGGAACGGAAGGCTCTGGAGGGCATTA
TCAGCAACTTTGGGCAGACTCCCTGTGAGTGTGTAAGGAGCCACATCCAACCTCGGCTCTCA
GCTGAGGAAGCAGCCCATCGCCTTGACGCCTGGACACTAACTCACCTAGCATCTTCCAGCA
CCTGGACGAACTCAAGGCATTCTTCGAGAGGTGACTGTGAGTGCCAGTGGGCTGCTGGGCA
CCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTACTTCAGCTTCAGCAAAGACCCC
ACCATGGGCAGCCACAAGACGCAGCGACTGCTGAGTGGCCCGTGGGTGCCAGGCAGTGGTGT
GAGTGGACAAGCACTGGCAGTGGCCCCGGATGGAAGCTGCTATTACGCGGTGGCCACTGGG
ATGGCAGCCTGCGGGTGAAGTCACTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGCTGCCAC
CTTGATGTAGTAACCTGCCTTGCACTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCG
GGACACCACGTGCATGGTGTGGCGGCTCCTGCATCAGGGTGGTCTGTGAGTAGGCCTGGCAC
CAAAGCCTGTGCAGGTCTGTATGGGCATGGGGCTGCAGTGAGCTGTGTGGCCATCAGCACT
GAAGTTGACATGGCTGTGTCTGGATCTGAGGATGGAAGTGTGATCATAACACTGTACGCCG
CGGACAGTTTGTAGCGGCCTACGGCCTCTGGGTGCCACATTCCCTGGACCTATTTCCACC
TGGCATTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGCGTGGGAACGTCTGGGGCC
CAGGTACCTACTCCTTGACCTGTATTCACTCAATGGGAAGTTGCGGGCTTCACTGGCCCT
GGCAGAGCAGCCTACAGCCCTGACGGTGACAGAGGACTTTGTGTGCTGGGCACCGCCAGT
GCGCCCTGCACATCCTCCAACATAACACTGCTCCCGGCCGCGCCTCCCTTGCCCATGAAG
GTGGCCATCCGCAGCGTGGCCGTGACCAAGGAGCGCAGCCACGTGCTGGTGGGCCTGGAGGA
TGGAAGCTCATCGTGGTGGTGGCGGGGAGCCCTCTGAGGTGCGCAGCAGCCAGTTCGCGC
GGAAGCTGTGGCGGTCTCGCGGGCGCATCTCCAGGTGTCTCGGGAGAGACGGAATACAAC
CCTACTGAGGCGCGCTGAACCTGGCCAGTCCGGCTGCTCGGGCCCCGCCCCGGCAGGCCTG
GCCCGGGAGGCCCCCGCCAGAAGTCGGCGGGAACACCCCGGGGTGGGCAGCCAGGGGGTGA
GCGGGGGCCACCCTGCCAGCTCAGGGATTGGCGGGCGATGTTACCCCTCAGGGATTGGCG
GGCGGAAGTCCCGCCCCCTCGCCGGCTGAGGGGCGCCCTGAGGGCCAGCACTGGCGTCT

23/310

FIGURE 23

MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL
RYTAVLKQQATQHSMALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVPN
HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELALETP
MEAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFYDGSTERVETE EGIGYDFRRP
LAQLREVHLRRFNLRRSALELFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQPGPIPPHTQV
RNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGR TYNDL
SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSKPIGVVNPKHAQLVREKYESFEDPAGTIDKFH
YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFD CSDRQFHSVAAAWQARLESPADV KELIP
EFFYFPDFLENQNGFDLGCLQLTNEKVGDVVLPPWASSPEDFIQQRQALESEYVSAHLHEW
IDLI FG YKQRGPA AEEALNVFY YCTYEGAVDL DHVTDERERKALEGII SNFGQTPCQLLKEP
HPTRL SAE EAAHRLARLD TNSPSIFQHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF
SFSKDPTMGSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWDGSLRV TALPRGKLL
SQLSCHLDVVTCLALDTCGIY LISGSRDTCMVWRLLHQGGLSVGLAPKPVQVLYGHGA AVS
CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPIFHLALGSEGQIVVQSSA
WERPGAQVTYSLHLYSVNGKLRASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA
PPLPMKVAIRSVAVTKERSHVLVGLEDGKLIVV VAGQPSEVRSSQFARKLWRSSRRISQVSS
GETEYNPTAR

FIGURE 24

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCC
CACGGCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCAT
CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTCTTC
TGGACCTTAACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTT
CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCC
GCACACTCCGTTACCACACTGGGTCAATTGGCATTGAGGCCCTCATCCTGACCCTTGTGCAG
ATAGCCCGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCTGTAGC
CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC
TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA
AATGCGTTCATGCTACTCATGCGAAACATTGTGAGGGTGGTCGTCCTGGACAAAGTCACAGA
CCTGCTGCTGTTCTTTGGGAAGCTGCTGGTGGTCGGAGGCGTGGGGGTCTGTCTTCTTTT
TTTTCTCCGGTTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCCACCTCAACTATTAC
TGGCTGCCCATCATGACCTCCATCCTGGGGGCTATGTATCGCCAGCGGCTTCTTCAGCGT
TTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACG
GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC
GAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGAACAGCTCCGGCCCTGATCCAGGACTGC
ACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGT
AAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG
AGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC
GTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTATCCCAGCTAC
TCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGAGA
TCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAACAA
AAAGATTTTATTAAAGATATTTTGTTAACCTC

FIGURE 25

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLF
WTLNWVLALGQCVLGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQ
IARVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAK
NAFMLLMRNIVRVVLDKVTDLLLFFGKLLVVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYY
WLPIMTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYSKSLKILGKKN
EAPPDNKKRKK

FIGURE 26

GAGTCTTGACCGCCGCCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGT
GGCTATGTTTCGTGTCCGATTTCCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCC
TTCTCTTCGTGGCCTCGGACGTGGATGCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTTC
CAGTGTGACCACGTGCAATATACGCTGGTTCAGTTTCTGGGTGGCAAGAACTTGAAACTGC
ATTTCTTGAGCATAAAGAACAGTTTCATTATTTTATTCTCATAACTGTGGAGCTAATGTAG
ACCTATTGGATATTCTTCAACCTGATGAAGACACTATATTCTTTGTGTGTGACTCCCATAGG
CCAGTCAATGTCGTCAATGTATACAACGATACCCAGATCAAATTACTCATTAAACAAGATGA
TGACCTTGAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATT
CAGGAAATGACAGTGATGGGTGAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATA
GTGGAGCAAACCATGCGGAGGAGGCAGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCT
CTTTGACTACGAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGG
CTTGATGCTGTCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGAC
CAGTGGGTGCAAGACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCTGTCAGCG
CCACGTTTCCCGCCACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCA
CACGGATCTCCTTTGAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGAC
AGCCTGTGCAACACCAGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAA
GCGGCTCCAGGAGTTCCTTGACAGATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCC
AGGCCATGGACATCTCCTTGAAGGAGAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAA
TTTGGGATGAAGGACATGCGCGTGCAGACTTTCAGCATTCATTTTGGGTTCAAGCACAAGTT
TCTGGCCAGCGACGTGGTCTTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCT
CAGGGACAGATCACTTCATCCAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTG
TACCATGGCCTGGAACCTCGCCAAGAAGCAGCTGCGAGCCACCAGCAGACCATTGCCAGCTGC
CTTTGCACCAACCTCGTCATCTCCCAGGGGCCTTTCCTGTACTGCTCTCTCATGGAGGGCAC
TCCAGATGTCATGCTGTTCTCTAGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCA
AGTCCTTTGTGTGTTGACAAAGAACCGGCGCTGCAAACCTGCTGCCCCCTGGTGATGGCTGCC
CCCCTGAGCATGGAGCATGGCACAGTGACCGTGGTGGGCATCCCCCAGAGACCGACAGCTC
GGACAGGAAGAACTTTTTTGGGAGGGCGTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGA
TGCTGCACAACCATTTTGACCTCTCAGTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTT
CTGGACGCACTTATTTCCCTCCTGTCCTAGGAAATTTGATTCTTCCAGAATGACCTTCTTATT
TATGTAACCTGGCTTTCATTTAGATTGTAAGTTATGGACATGATTTGAGATGTAGAAGCCATT
TTTTATTAAATAAAATGCTTATTTTAGGAAA

FIGURE 27

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETAF
LEHKEQFHYFILINCGANVDLLDILQPDEDTIFFVCDSHRPVNVVNVYNDTQIKLLIKQDDD
LEVPAIEDIFRDEEEDEEHSGNDSGSEPSEKRTRELEEIVEQTMRRRQRREWEARRRDILF
DYEQYEHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGVLQRH
VSRHNRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSL CNTSYTAARFKLWSVHGQKR
LQEF LADMGLPLKQVKQKFQAMD ISLKENLREMI EESANKFGMKDMRVQTFSIHFGFKHKFL
ASDVVFATMSLMESPEKDGSGTDHFIQALDSLRSNLDKLYHGLELAKKQLRATQQTIASCL
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKSFVCSTKNRRCKLLPLVMAAP
LSMEHGTVTTVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNFDSLVI ELKAEDRSKFL
DALISLLS

FIGURE 28

GTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCGTGGCTATGNTCGTGTCCGATTTCCGCA
AAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCCTTCTCTTCGTGGCCTCGGANGTGGAT
GCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTCCAGTGTGACCANGTGCAATATANGCT
GGTTCAGTTTCTGGGTGGCAAGAACTTGAACTGCATTTCTTGAGCATAAAGAACAGTTTC
ATTATTTTATTCTCATAAACTGTGGAGCTAATGTAGACCTATTGGATATTCTTCAACCTGAT
GAAGACACTATATTCTTTGTGTGTGACACCCATAGGCCAGTCAATGTTGTCAATGTATACAA
CGATACCC

-29/3-10

FIGURE 29

CAGGAACCCTCTCTTTGGGTCTGGATTGGGACCCCTTTCCAGTACCATTTTTTCTAGTGAAC
CACGAAGGGACGATACCAGAAAAACACCCTCAACCCAAAGGAAATAGACTACAGCCCCAATTG
GCTGACTTTGGCTATAGAAAAAAGAAAGGAACGAAAAGAGACAGTTTTTTTTTGGAAAGCTAA
GTCTTCCCTTTATCGAGTCAAGAAACCCCCCTTCTTGAGCTATTTACAGCTTTTAACAATT
GAGTAAAGTACGCTCCGGTCACCATGGTGACAGCCGCCCTGGGTCCCGTCTGGGCAGCGCTC
CTGCTCTTTCTCCTGATGTGTGAGATCCGTATGGTGGAGCTCACCTTTGACAGAGCTGTGGC
CAGCGGCTGCCAACGGTGCTGTGACTCTGAGGACCCCTGGATCCTGCCCATGTATCCTCAG
CCTCTTCTCCTCCGGCCGCCCCACGCCCTGCCTGAGATCAGACCCTACATTAATATCACCATC
CTGAAGGGTGACAAAGGGGACCCAGGCCCAATGGGCCTGCCAGGGTACATGGGCAGGGAGGG
TCCCCAAGGGGAGCCTGGCCCTCAGGGCAGCAAGGGTGACAAGGGGGAGATGGGCAGCCCCG
GCGCCCCGTGCCAGAAGCGCTTCTTCGCCCTTCTCAGTGGGCCGCAAGACGGCCCTGCACAGC
GGCGAGGACTTCCAGACGCTGCTCTTCGAAAGGGTCTTTGTGAACCTTGATGGGTGCTTTGA
CATGGCGACCGGCCAGTTTGCTGCTCCCTGCGTGGCATCTACTTCTTCAGCCTCAATGTGC
ACAGCTGGAATTACAAGGAGACGTACGTGCACATTATGCATAACCAGAAAGAGGCTGTATC
CTGTACGCGCAGCCCAGCGAGCGCAGCATCATGCAGAGCCAGAGTGTGATGCTGGACCTGGC
CTACGGGGACCGCGTCTGGGTGCGGCTCTTCAAGCGCCAGCGCGAGAACGCCATCTACAGCA
ACGACTTCGACACCTACATCACCTTCAGCGGCCACCTCATCAAGGCCGAGGACGACTGAGGG
CCTCTGGGCCACCCTCCCGGCTGGAGAGCTCAGGTGCTGGTCCCGTCCCTGCAGGGCTCAG
TTTSCACTGCTGTGAAGCAGGAAGGCCAGGGAGGTCCCCGGGGACCTGGCATTCTGGGGAGA
CCCTGCTTCTATCTTGGCTGCCATCATCCCTCCAGCCTATTTCTGCTCCTCTCTCTCTCT
TGGACCTATTTTAAGAAGCTTGCTAACCTAAATATTCTAGAACTTTCCAGCCTCGTAGCCC
AGCACTTCTCAAACCTTGGAATGCATGCGAATCACCCGGGGTTCGTGTTAAATGCAGATTCT
GACTCAGCAGGTCTGAGTGGGTCCAGGATTCTGTGTTTCTCATATGTTCTGGGTGATGCTG
ATGGGGTCAGTCTATGAACCACACTGGAGCAACCAGGTTCTAGGACTTTCTCAATATTCTAG
TACTTTCTGAACATTCTGGAATCCTCCCCACATTCTAGAATTCTCCCAACATTTTTTTTTCT
TGAGACAGAGTCTTGCTCTGTTGCCAGGCTAGAGTGCAGTGGTGCAATCTCAGTTCAGTGC
AACCTCTGCCTCCCGGGTCAAGCGATTCTTCTGCCTCAGCCTCCCTAGTGGCTGGGATTAC
AGGCGCCTGCTACCATGCCTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCATA
TTGGCCAGGCTGGTCTTGAACCTCTGACTTCAGGTGACCCACCCGCCTCGGCCTCTCAAAAT
GCTGGGATTACAGGTGTGAGCCACCGTGCCCTGGCCAATTCCAACATTCTTAAATTCTCTCAT
CCCTCCAGGGCTCCCGTGCTATGTTCTCTTTACCCCTTCCCCCTCTTCTCTTGCTCAGGCC
TGCACCACTGCAGCCACCGTTCATTTATTCAATTAACACTGAGCACTCACTCTGTGCT
GGGTCCCGGGAAGGGTGAGGGGGTCAGACACAGGCCCTGCCCTGCCCTCAGTGACTGGCCA
GTCCAGCCCAGGCGGGGAGAGATGTGTACATAGTTTTTAAAGCAGACCCAGAGCTCATGGGG
GCCTGTGTTCTGGGTGTTCAAGTGCTGCTGGTCTCCATTACCCACTGCTCCCCAAGGCTGG
TGGGACGGGGTCCCGGTGGCAGGGGCAGGTATCTCCTTCCCGTTCCTCATCCACCTGCCCAG
TGCTCATCGTTACAGCAAACCCAGGGGGCTTGGCCAGGTCAAGGGTTCGTGTGAGGAGAGG
ACCCAGGAGTGTGGGGGCATTTGGGGGGTGAAGTGGCCCCCGAAGAATGGAACCCACACCCA
TAGCTCTCCCCACAGCTGATACGGCATCCTGCGAGAAGACCTGCCCTCCTCACTGGGATCCC
CTTCTGCTCCTCCTCCAGGGCTCTGCCAGGGCCTTGCTCAGTCCCTTCCACCAAAGTCATCT
GAACTTCCGTTTCCCCAGGGCCTCCAGCTGCCCTCAGACACTGATGTCTGTCCCCAGGTGCT
CTCTGCCCCCTCATGCCCTCTCACCGGCCAGTGCCCCGACTCTCCAGGCTTTATCAAGGTG
CTAAGGCCCCGGGTGGGCAGCTCCTCGTCTCAGAGCCCTCCTCCGGCCTGGTGCTGCCTTTAC
AAACACCTGCAGGAGAAGGGGCCACGGAAGCCCCAGGCTTTAGAGCCCTCAGCAGGTCTGGGG
AGCTAGAGCAAAGGAGGGACCTCAGGCCTTCCGTTTCTTCTTCCAGGGTGGGGTGGCCTGGT
GTTCCCTAGCCTTCCAAACCCAGGTGGCCTGCCCTTCTCCCCAGAGGGAGGGCGGCCTCCGC
CCATTGGTGCTCATGCAGACTCTGGGGCTGAGGTGCCCCGGGGGGTGATCTCTGGTGCTCAC
AGCCGAGGGAGCCGTGGCTCCATGGCCAGATGACGGAAACAGGGTCTGACCAAGTGCCAGGA
AGACCTGTGCTATAAACCACCTGCTGATCCTGCCCTGCCTGACCCGCCACGCCCTGCC
GTCCAGCATGATTAAAGAATGCTGTCTCCTCTTGAAAAA

30/310

FIGURE 30

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSSGRPH
ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPPQGSKGDKGEMGSPGAPCQKRF
FAFSVGRKTALHSGEDFQTLLFERVVFVNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET
YVHIMHNQKEAVILYAQPSEERSIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYIT
FSGHLIKAEDD

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-75

Clq domain proteins.

amino acids 144-178, 78-111 and 84-117

31/310

FIGURE 31

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCGACCCGCCAGGAAAGACTG
AGGCCGCGGCCTGCCCCGCGGCTCCCTGCGCCGCGCGCCTCCCGGGACAGAAGATGTG
CTCCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGG
GCTGCCCATCCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGG
ACCACGGTGCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCAT
CACCATGCTCGACGCAGGCAGCTTTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTCAC
AGAACCAGATCGCCAGCCTGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTG
GACCTGACGGCCAACAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCT
CGAGCGCCTCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGC
TCGACCGCCTCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCGCTGCGC
CTGCCCCGCTGCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCAT
CCTGGACACTGCCAACGTGGAGGCGCTGCGGCTGGCTGGTCTGGGGCTGCAGCAGCTGGAGC
AGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAG
CGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACGCGCCTGCGGCTGGCCGGCAACAC
CCGCATTGCCAGCTGCGGCCCGAGGACCTGGCCGGCCTGGCTGCCCTGCAGGAGCTGGATG
TGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCGCTGCGG
CTGCTGGCAGCTGCCCCGAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGGCCCCTG
GGTGCGCGAGAGCCACGTCACTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCGCCCA
AGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAGCCACCACC
ACCACAGCCACAGTGCCACCACGAGGCCCGTGGTGCGGGAGCCACAGCCTTGTCTTCTAG
CTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCCTCCA
CTGCCCCACCGACTGTAGGGCCTGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTC
AATGGGGGCACATGCCACCTGGGGACACGGCACCACTGGCGTGCTTGTGCCCCGAAGGCTT
CACGGGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCCAGCCCTACACCAGTCA
CGCCGAGGCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGC
GTGGGGCTGCAGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTA
TCGCAACCTATCGGGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTG
AGTACACGGTCACCCAGCTGCGGCCCAACGCCACTTACTCCGTCTGTGTCATGCCTTTGGGG
CCCGGGCGGGTGCCGGAGGGCGAGGAGGCCTGCGGGGAGGCCATACACCCCGAGCCGTCCA
CTCCAACCACGCCCCAGTCACCCAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCG
CCCTGGCCGCGGTGCTCCTGGCCGCGCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGG
GGGCGGGCCATGGCAGCAGCGGCTCAGGACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCCCT
GGAACCTGGAGGGAGTGAAGGTCCCCTTGAGGCCAGGCCCGAAGGCAACAGAGGGCGGTGGAG
AGGCCCTGCCCAGCGGGTCTGAGTGTGAGGTGCCACTCATGGGCTTCCCAGGGCCTGGCCTC
CAGTCACCCCTCCACGCAAAGCCCTACATCTAAGCCAGAGAGAGACAGGGCAGCTGGGGCCG
GGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCCACACCACGTAAGTTCTCAGTCC
CAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCTGGGCCCTGTTCCCTCTGGA
CCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCCCTAACGTCCCCAGAAC
CGAGTGCCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTCCCTGGGCACGGCG
GGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCCACTCCAGGCGGA
CCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTG
TGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAGAACTGGAAAGGAAGATGCTTTA
GGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCTG
GGAAGATGTTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTTGTAAGACAAACGATGATATG
AAGGCCTTTTGTAAGAAAAAATAAAAGATGAAGTGTGAAA

32 / 310

FIGURE 32

MCSRVP L L L L L L L L L L A L G P G V Q G C P S G C Q C S Q P Q T V F C T A R Q G T T V P R D V P P D T V G L Y V F E N
G I T M L D A G S F A G L P G L Q L L D L S Q N Q I A S L P S G V F Q P L A N L S N L D L T A N R L H E I T N E T F R G L R
R L E R L Y L G K N R I R H I Q P G A F D T L D R L L E L K L Q D N E L R A L P P L R L P R L L L L D L S H N S L L A L E P
G I L D T A N V E A L R L A G L G L Q Q L D E G L F S R L R N L H D L D V S D N Q L E R V P P V I R G L R G L T R L R L A G
N T R I A Q L R P E D L A G L A A L Q E L D V S N L S L Q A L P G D L S G L F P R L R L L A A A R N P F N C V C P L S W F G
P W V R E S H V T L A S P E E T R C H F P P K N A G R L L L E L D Y A D F G C P A T T T T A T V P T T R P V V R E P T A L S
S S L A P T W L S P T A P A T E A P S P P S T A P P T V G P V P Q P Q D C P P S T C L N G G T C H L G T R H H L A C L C P E
G F T G L Y C E S Q M G Q G T R P S P T P V T P R P P R S L T L G I E P V S P T S L R V G L Q R Y L Q G S S V Q L R S L R L
T Y R N L S G P D K R L V T L R L P A S L A E Y T V T Q L R P N A T Y S V C V M P L G P G R V P E G E E A C G E A H T P P A
V H S N H A P V T Q A R E G N L P L L I A P A L A A V L L A A L A A V G A A Y C V R R G R A M A A A A Q D K G Q V G P G A G
P L E L E G V K V P L E P G P K A T E G G G E A L P S G S E C E V P L M G F P G P G L Q S P L H A K P Y I

33/310

FIGURE 33

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCAT
CTTCATCATTTCATATGAGGAAATAAGTGGTAAAATCCTTGGAATACAATGAGACTCATCAG
AAACATTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAG
AAGAAAGGGAACTGATGACCAACTGCTCCAACATGTCTCTAAGAAAGGTTCCCGCAGACTTG
ACCCAGCCACAACGACACTGGATTATCCTATAACCTCCTTTTTCAACTCCAGAGTTCAGA
TTTTCATTTCTGTCTCCAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCAACAGCTGG
ATCTCAAAACCTTTGAATTCAACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTG
AAGAGTGTAACCTTGGTATTTACTGGCAGGTCTCAGGTATTTAGATCTTTCTTTAATGACTT
TGACACCATGCCATCTGTGAGGAAGCTGGCAACATGTCACACCTGGAAATCCTAGGTTTGA
GTGGGGCAAAAATACAAAATCAGATTTCCAGAAAATTGCTCATCTGCATCTAAATACTGTC
TTCTTAGGATTCAGAACTCTTCCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAAC
AAAACCTGCACATTGTTTTACCAATGGACACAAATTTCTGGGTTCTTTGCGTGATGGAATCA
AGACTTCAAAAATATTAGAAATGACAAATATAGATGGCAAAAGCCAATTTGTAAGTTATGAA
ATGCAACGAAATCTTAGTTTAGAAAAATGCTAAGACATCGGTTCTATTGCTTAATAAGTTGA
TTTACTCTGGGACGACCTTTTCCTTATCTTACAATTTGTTTGGCATAACATCAGTGGAACT
TTCAGATCCGAAATGTGACTTTTGGTGGTAAGGCTTATCTTGACCACAATTCATTGACTAC
TCAAACTACTGTAATGAGAACTATAAAATTGGAGCATGTACATTTTCCAGAGTGTTTACATTCA
ACAGGATAAAATCTATTGCTTTTGACCAAAATGGACATAGAAAACCTGACAATATCAAATG
CACAAATGCCACATGCTTTTCCGAATTAATCTACGAAATTCGAATATTTAAATTTTGCC
AATAATATCTTAACAGACGAGTTGTTTAAAGAACTATCCAAGTGCCTCACTTGAAAATCTCT
CATTTTGAATTGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACAC
CCTTGGAACACTTGGATCTGAGTCAAATCTATTACAACATAAAAATGATGAAAATTGCTCA
TGGCCAGAACTGTGGTCAATATGAATCTGTCTACAATAAATTGTCTGATTCTGTCTTCAG
GTGCTTGCCCAAAAGTATTCAAATACTTGACCTAAATAATAACCAATCCAACTGTACCTA
AAGAGACTATTCATCTGATGGCCTTACGAGAACTAAATATTGCATTTAATTTTCTAACTGAT
CTCCCTGGATGCAGTCATTTTCCAGTAGACTTTTCCAGTTCTGAACATTGAAATGAACCTTCATTCT
CAGCCCATCTCTGGATTTTGTTCAGAGCTGCCAGGAAGTTAAACTCTAAATGCGGGAAGAA
ATCCATTCCGGTGTACCTGTGAATTAATAAATTTTCAATTCAGCTTGAAACATATTCCAGAGGTC
ATGATGGTTGGATGGTCAGATTCATACACCTGTGAATACCCTTTAAACCTAAGGGGAACCTAG
GTTAAAGACGTTTCATCTCCACGAATTAATCTTGCAACACAGCTCTGTTGATTGTCCACATTG
TGGTTATTATGCTAGTTCTGGGTTGGCTGTGGCCTTCTGCTGTCTCCACTTTGATCTGCCC
TGGTATCTCAGGATGCTAGGTCAATGCACACAAACATGGCACAGGGTTAGGAAAACAACCCA
AGAACAACCTCAAGAGAAATGTCCGATTCCACGCATTTATTTTATACAGTGAACATGATTCTC
TGTGGGTGAAGAATGAATTGATCCCCAATCTAGAGAAGGAAGATGGTTCTATCTTGATTGTC
CTTTATGAAAGCTACTTTGACCCTGGCAAAAGCATTAGTGAAAATATTGTAAGCTTCATTGA
GAAAAGCTATAAGTCCATCTTTGTTTGTCTCCCACTTTGTCCAGAATGAGTGGTGCCATT
ATGAATTCTACTTTGCCCAACCAATCTCTTCCATGAAAATTCTGATCATATAATTCTTATC
TTACTGGAACCCATTCCATTCTATTGCATTCCCACCAGGTATCATAACTGAAAGCTCTCCT
GGAAAAAAAGCATACTTGGAATGGCCCAAGGATAGGCGTAAATGTGGGCTTTTCTGGGCAA
ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCAGAGAAATGTATGAACTGCAGACA
TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCTCTCTGATGAGAACAGATTGTCT
ATAAAATCCCACAGTCCTTGGGAAGTTGGGGACCACATACACTGTTGGGATGTACATTGATA
CAACCTTTATGATGGCAATTTGACAAATTTTATTAATAAATAAATAAGTTATTCCCTTCATA
TCAGTTTCTAGAAGGATTTCTAAGAATGTATCCTATAGAAACACCTTCACAAGTTTATAAGG
GCTTATGGAAAAAGGTGTTTCATCCCAGGATTGTTTATAATCATGAAAAATGTGGCCAGGTGC
AGTGGCTCACTCTTGTAATCCCAGCACTATGGGAGGCCAAGGTGGGTGACCCACGAGGTCAA
GAGATGGAGACCATCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAATTA
GCTGGGCGTGATGGTGCACGCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCG
CTTGAACCCGGGAGGTGGCAGTTGCAGTGAGCTGAGATCGAGCCACTGCACTCCAGCCTGGT
GACAGAGCGAGACTCCATCTCAAAAAAAGAAAAAAGAAAAAATGGAACATCC
TCATGGCCACAAAATAAGGTCTAATTCAATAAATTATAGTACATTAATGTAATATAATATTA
CATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCTGGTATGGAAAAACATATTAATAT
GTTATAAACTATTAGGTTGGTGCAAACTAATTGTGGTTTTTGGCATTGAAATGGCATTGAA
ATAAAAGTGTAAGAAATCTATACCAGATGTAGTAACAGTGGTTTGGGTCTGGGAGGTGGA
TTACAGGGAGCATTTGATTTCTATGTTGTGTATTTCTATAATGTTTGAATTGTTTAGAATGA
ATCTGTATTTCTTTTATAAGTAGAAAAAATAAAGATAGTTTTTACAGCCT

34/340

FIGURE 34

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQ
LQSSDFHSVSKLRVLILCHNRIQQDLKTFEFNKELRYLDLSNNRLKSVTWYLLAGLRYLDL
SFNDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLP
ILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLL
LNKVDLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFR
VFYIQQDKIYLLLTKMDIENLTISNAQMPHMLFPNYPTKFQYLNFANNILTDELFKRTIQLP
HLKTLILNGNKLETLSLVSCFANNTPLEHLDLSQNLLQHKNDENC SWPETVVMNLSYNKLS
DSVFRCLPKSIQILD LNNNQIQTVPKETIHLMALRELNIAFNFLTDLPGCSHF S RLSVLNIE
MNFILSPSLDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETYSEVMMVGWSDSYTCEYPLN
LRGTRLKDVHLHELSCNTALLIVTIVVIMLVGLAVAFCC LHFDPWYLRMLGQCTQTWHRV
RKTTQEQLKRNVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENI
VSFIEKSYKSIFVLSPNFVQNEWCHYEFYFAHNLFHENS DHIILILLEPIPFYCIPTRYHK
LKALLEKKAYLEWPKDRRKCGLFWANLRAAINVNVLATREMYELQTFTELNEESRGSTISLM
RTDCL

35/310

FIGURE 35A

GGGGGCTTTCTTGGGCTTGGCTGCTTGGAAACACCTGCCTCCAAGGACCGGCCTCGGAGGGGT
CGCCGGGAAAGGGAGGGAAGAAGGAAGGGCGGGGCCCGCCCTGCGCCCGCCCGCGCCT
CTGCGCGCCCTGTCCGCCCCGGCCCCAGCCCAGCCCAGCCCCGCGGGCCGGTCACACGCGCA
GCCAGCCGGCCGCCTCCCGCGCCCCAAGCGCGCCGCTCTGCTGTGCCCTGCGCCCTTGCCCCG
CGCCAGCTTCTGCGCCCGCAGCCCGCCCGGCGCCCCCGGTGACCGTGACCCTGCCCTGGGCG
CGGGGCGGAGCAGGCATGTCGCCCGGGGACCGCTACCCAGCGCTGGCCCTGGTGCTCCT
GGCAGTGACCCTGGCCGGGGTTCGGAGCCCAGGGCGCAGCCCTCGAGGACCCTGATTATTACG
GGCAGGAGATCTGGAGCCGGGAGCCCTACTACGCGCGCCCGGAGCCCGAGCTCGAGACCTTC
TCTCCGCGCTGCCTGCGGGGCCCGGGAGGAGTGGGAGCGGCGCCCGCAGGAGCCCAGGCC
GCCCAAGAGGGGCCACCAAGCCCAAGAAAGCTCCCAAGAGGGAGAAGTCGGCTCCGGAGCCGC
CTCCACCAGGTAAACACAGCAACAAAAAGTTATGAGAACCAAGAGCTCTGAGAAGGCTGCC
AACGATGATCACAGTGTCGTGTGGCCCGTGAAGATGTCAGAGAGAGTTGCCACCTCTTGG
TCTGGAAACCTTAAAAATCACAGACTTCCAGCTCCATGCCTCCACGGTGAAGCGCTATGGCC
TGGGGGCACATCGAGGGAGACTCAACATCCAGGCGGGCATTAAATGAAAATGATTTTTATGAC
GGAGCGTGGTGCGCGGGAAGAAATGACCTCCAGCAGTGGATTGAAGTGGATGCTCGGCGCCT
GACCAGATTCACTGGTGTCACTCAAGGGAGGAACTCCCTCTGGCTGAGTGACTGGGTGA
CATCCTATAAGGTCATGGTGAGCAATGACAGCCACACGTGGGTCACTGTAAAGATGGATCT
GGAGACATGATATTTGAGGGAAACAGTGAGAAGGAGATCCCTGTTCTCAATGAGCTACCCGT
CCCCATGGTGGCCCGCTACATCCGCATAAACCCCTCAGTCCTGGTTTGATAATGGGAGCATCT
GCATGAGAATGGAGATCCTGGGCTGCCCCTGCCAGATCCTAATAATTATTATCACCGCCGG
AACGAGATGACCACCACTGATGACCTGGATTTAAGCACCACAATTATAAGGAAATGCGCCA
GTTGATGAAAGTTGTGAATGAAATGTGTCCCAATATCACCAGAATTTACAACATTGGAAAAA
GCCACCAGGGCCTGAAGCTGTATGCTGTGGAGATCTCAGATCACCTGGGGAGCATGAAGTC
GGTGAGCCCGAGTTCCACTACATCGCGGGGGCCACGGCAATGAGGTGCTGGGCCGGGAGCT
GCTGCTGCTGCTGGTGCAGTTCGTGTGTGAGGAGTACTTGGCCCGGAATGCGCGCATCGTCC
ACCTGGTGGAGGAGACGCGGATTACAGTCCTCCCTCCCTCAACCCCGATGGCTACGAGAAG
GCCTACGAAGGGGGCTCGGAGCTGGGAGGCTGGTCCCTGGGACGCTGGACCCACGATGGAAT
TGACATCAACAACAACCTTTCTGATTTAAACACGCTGCTCTGGGAGGCAGAGGATCGACAGA
ATGTCCCCAGGAAAGTTCCCAATCACTATATTGCAATCCCTGAGTGGTTTCTGTGCGAAAAAT
GCCACGGTGGCTGCCGAGACCAGAGCAGTCATAGCCTGGATGGAAAAAATCCCTTTTGTGCT
GGGCGGCAACCTGCAGGGCGGCGAGCTGGTGGTGGCGTATCCCTACGACCTGGTGGGTCCC
CCTGGAAGACGCAGGAACACACCCCCACCCCCGATGACCACGTGTTCCGCTGGCTGGCCTAC
TCCTATGCCTCCACACACCGCCTCATGACAGACGCCCCGAGGAGGGTGTGCCACACGGAGGA
CTTCAGAAAGGAGGAGGGCACTGTCAATGGGGCCTCCTGGCACACCGTCGCTGGAAGTCTGA
ACGATTTTCACTACCTTCATACAACTGCTTCGAACTGTCCATCTACGTGGGCTGTGATAAA
TACCCACATGAGAGCCAGCTGCCCCGAGGAGTGGGAGAATAACCGGGAATCTCTGATCGTGTT
CATGGAGCAGGTTTCATCGTGGCATTAAGGCTTGGTGAGAGATTCATGGAAGGAATCC
CAAACGCCATTATCTCCGTAGAAGGCATTAACCATGACATCCGAACAGCCAACGATGGGGAT
TACTGGCGCCTCCTGAACCTGGAGAGTATGTGGTCACAGCAAAGGCCGAAGGTTTCACTGC
ATCCACCAAGAACTGTATGGTTGGCTATGACATGGGGGCCACAAGGTGTGACTTCACACTTA
GCAAAACCAACATGGCCAGGATCCGAGAGATCATGGAGAAGTTTGGGAAGCAGCCCGTCAGC
CTGCCAGCCAGGCGGCTGAAGCTGCGGGGGCGGAAGAGACGACAGCGTGGGTGACCCCTCTG
GGCCCTTGAGACTCGTCTGGGACCATGCAAATTAACCAACCTGGTAGTAGCTCCATAGTG
GACTCACTCACTGTTGTTTCTCTGTAATTCAAGAAGTGCCTGGAAGAGAGGGTGCATTGTG
AGGCAGGTCCCAAAAGGGAAGGCTGGAGGCTGAGGCTGTTTTCTTTCTTTGTTCCCATTTA
TCCAAATAACTTGGACAGAGCAGCAGAGAAAAGCTGATGGGAGTGAGAGAACTCAGCAAGCC
AACCTGGGAATCAGAGAGAGAAGGAGAAGGAGGGGAGCCTGTCCGTTCAAGACCTCTGGCTGC

36/310
FIGURE 35B

ATAGAAAAGGATTCTGGTGCTTCCCCTGTTTGCGTGCGCAGCAAGGGTTCACGTGCATTTGC
AATTTGCACAGCTAAAATTGCAGCATTTCCCCAGCTGGGCTGTCCCAAATGTTACCATTGA
GATGCTCCCAGGCGTCCTAAGAGAATCCACCCTCTCTGGCCCTGGGACATTGCAAGCTGCTA
CAAATAAATTCTGTGTTCTTTTGACAATAGCGTCATTGCCAAGTGCACATCAGTGAGCCTCT
TGAATCTGTTTAGTCTCCTTTTTCAACAAAGGAGTGTGTTCAGAAAAGGAGAGAGAGGCTGA
GATCATTGAGGAGTTTGTGTTGGGCAGCAAGCATGGAGCTTCTTGACAAAATTCTGGGTCCATA
AACAACCCCCAAAGTCCCTGCTGATCCAGTAGCCCTGGAGGTCCCCAGGTAGGGAGAGCCA
GAGGTGCCAGCCTTCCTGAAGGGCCAGAAAATTTAGCCTGGATCTCCTCTTTTACCTGCTAG
GACTGGAAAGAGCCAGAAGTGGGGTGGCCTGAAGCCCTCTCTCTGCTTGAGGTATTGCCCT
GTGTGGAATTGAGTGCTCATGGGTGGCCTCATATCAGCCTGGGAGTTATTTTATGATATGTA
GAATGCCAGATCTTCCAGATTAGGCTAAATGTAATGAAAACCTCTTAGGATTATCTGTGGAG
CATCAGTTTGGGAAGAATTATTGAATTATCTTGCAAGAAAAAAGTATGTCTCACTTTTTGTT
AATGTTGCTGCCTCATTGACCTGGGAAAAATGAAAAAAAAAAATAAGCAAATGGTAAGACC
CTTAAA

37/310
FIGURE 36

MSRPGTATPALALVLLAVTLAGVGAQGALEDPDYYGQEIWSREPYARPEPELETFSPPLP
AGPGEEWERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS
VRVAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCA
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSTWVTVKNGSGDMIF
EGNSEKEIPVLNELPVPMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTT
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEF
HYIAGAHGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGG
SELGGWSLGRWTHDGDIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAA
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSAST
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHES
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRLL
NPGEYVVTAKAEGFTASTKNCMVGYDMGATRCDFTLSTNMARI REIMEKFGKQPVSLPARR
LKLGRGRRRQRG

38 /310
FIGURE 37

CTAAGAGGACAAGATGAGGCCCGGCTCTCATTTCTCCTAGCCCTTCTGTTCTTCCTTGGCC
AAGCTGCAGGGGATTTGGGGGATGTGGGACCTCCAATTCCCAGCCCCGGCTTCAGCTCTTTC
CCAGGTGTTGACTCCAGCTCCAGCTTCAGCTCCAGCTCCAGGTCCGGGCTCCAGCTCCAGCCG
CAGCTTAGGCAGCGGAGGTTCTGTGTCCAGTTGTTTTCCAATTTACCCGGCTCCGTGGATG
ACCGTGGGACCTGCCAGTGCTCTGTTCCCTGCCAGACACCACCTTTCCCGTGGACAGAGTG
GAACGCTTGGAATTCACAGCTCATGTTCTTTCTCAGAAGTTTGAGAAAGAACTTTCTAAAGT
GAGGGAATATGTCCAATTAATTAGTGTGTATGAAAAGAAACTGTTAAACCTAACTGTCCGAA
TTGACATCATGGAGAAGGATAACCATTTCTTACACTGAACTGGACTTCGAGCTGATCAAGGTA
GAAGTGAAGGAGATGGAAAACTGGTCATACAGCTGAAGGAGAGTTTTTGGTGGAAGCTCAGA
AATTGTTGACCAGCTGGAGGTGGAGATAAGAAATATGACTCTCTTGGTAGAGAAGCTTGAGA
CACTAGACAAAAACAATGTCCTTGCCATTGCGCCGAGAAATCGTGGCTCTGAAGACCAAGCTG
AAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCTGTCGTCCACCCTCCTCCCACTCCAGG
GAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGTGGTTCAGCTCAACTGGA
GAGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTCCCCAGCATCCAAACAAA
GGACTGTATTGGGTGGCGCCATTGAATACAGATGGGAGACTGTTGGAGTATTATAGACTGTA
CAACACACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTTGCGGATCACCTATGGCC
AAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAACATGTACAACACCGGGAAT
ATTGCCAGAGTTAACCTGACCACCAACACGATTGCTGTGACTCAAACCTCTCCCTAATGCTGC
CTATAATAACCGCTTTTTCATATGCTAATGTTGCTTGGCAAGATATTGACTTTGCTGTGGATG
AGAATGGATTGTGGGTTATTTATTCAACTGAAGCCAGCACTGGTAACATGGTGAATTAGTAAA
CTCAATGACACCACACTTCAGGTGCTAAACACTTGGTATACCAAGCAGTATAAACCATCTGC
TTCTAACGCCTTCATGGTATGTGGGGTTCTGTATGCCACCCGTAATGAAACACCAGAACAG
AAGAGATTTTTTACTATTATGACACAAACACAGGGAAAGAGGGCAAACCTAGACATTGTAATG
CATAAGATGCAGGAAAAAGTGCAGAGCATTAACTATAACCCTTTTGACCAGAACTTTATGT
CTATAACGATGGTTACCTTCTGAATTATGATCTTTCTGTCTTGCAGAAGCCCCAGTAAAGCTG
TTTAGGAGTTAGGGTGAAAGAGAAAAATGTTTGTTGAAAAAATAGTCTTCTCCACTTACTTAG
ATATCTGCAGGGGTGTCTAAAGTGTGTTTCAATTTTGCAAGCAATGTTTAGGTGCATAGTTCTA
CCACACTAGAGATCTAGGACATTTGTCTTGATTGGTGAGTTCTCTTGGGAATCATCTGCCT
CTTCAGGCGCATTTTGCAATAAAGTCTGTCTAGGGTGGGATTGTCAGAGGTCTAGGGGCACT
GTGGGCCTAGTGAAGCCTACTGTGAGGAGGCTTCACTAGAAGCCTTAAATTAGGAATTAAGG
AACTTAAACTCAGTATGGCGTCTAGGGATTCTTTGTACAGGAAATATTGCCCAATGACTAG
TCCTCATCCATGTAGCACCCTAATTCTTCCATGCCTGGAAGAAACCTGGGGACTTAGTTAG
GTAGATTAATATCTGGAGCTCCTCGAGGGACCAATCTCCAACCTTTTTTTTTCCCCTCACTAG
CACCTGGAATGATGCTTTGTATGTGGCAGATAAGTAAATTTGGCATGCTTATATATTCTACA
TCTGTAAAGTGCTGAGTTTTATGGAGAGAGGCCTTTTTATGCATTAAATTGTACATGGCAAA
TAAATCCCAGAAGGATCTGTAGATGAGGCACCTGCTTTTTCTTTCTCTCATTGTCCACCTT
ACTAAAAGTCAGTAGAATCTTCTACCTCATAACTTCCTTCAAAGGCAGCTCAGAAGATTAG
AACCAGACTTACTAACCAATTCACCCCCCAACACCCCTTCTACTGCCTACTTTAAAAAA
ATTAATAGTTTTCTATGGAACTGATCTAAGATTAGAAAAATTAATTTTCTTTAATTTTATTA
TGGACTTTTATTTACATGACTCTAAGACTATAAGAAAATCTGATGGCAGTGACAAAGTGCTA
GCATTTATTGTTATCTAATAAAGACCTTGGAGCATATGTGCAACTTATGAGTGTATCAGTTG
TTGCATGTAATTTTTGCCTTTGTTTAAAGCCTGGAACCTGTAAGAAAATGAAAATTTAATTTT
TTTTTCTAGGACGAGCTATAGAAAAGCTATTGAGAGTATCTAGTTAATCAGTGCAGTAGTTG
GAAACCTTGCTGGTGTATGTGATGTGCTTCTGTGCTTTTGAATGACTTTATCATCTAGTCTT
TGTCTATTTTTCTTTGATGTTCAAGTCCTAGTCTATAGGATTGGCAGTTTAAATGCTTTAC
TCCCCCTTTTAAATAAATGATTAAATGTGCTTTGAAAAAATAAAAAAAAAAAAAAAAAAAAAA

39 / 310

FIGURE 38

MRPGLSFLLALLFFLGQAAGDLGDVGPPIPSPGFSSFPGVDSSSSFSSSSRSGSSSSRSLGS
GGSVSQLFSNFTGSVDDRGTCCSVSLPDTTFPVDRVERLEFTAHVLSQKFELSKVREYV
QLISVYEKKLLNLTVRIDIMEKDTISYTELDFELIKVEVKEMEKLVIQLKESFGGSSEIVDQ
LEVEIRNMTLLVEKLETLDKNNVLAIRREIVALKTCLKECEASKDQNTPVVHPPPTPGSCGH
GGVNIISKPSVVQLNWRGFSYLYGAWGRDYSPQHNPNGLYWVAPLNTDGRLLLEYRLYNTLD
DLLLYINARELRITYGQSGTAVYNNNMYVNMVNTGNIARVNLTTNTIAVTQTLNAAAYNNR
FSYANVAWQDIDFAVDENGLWVIYSTEASTGNMVISKLNDTTLQVLNTWYTKQYKPSASNAF
MVCGLVLYATRTMNTRTEEIFYYYDTNTGKEGKLDIVMHKMQEKVQSINYNPFDQKLYVYNDG
YLLNYDLSVLQKPQ

40/310

FIGURE 39

GCTCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCTGTCGTCCAC
CCTCCTCCCACTCCAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGT
GGTTCAGCTCAACTGGAGAGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTC
CCCAGCATCCAAACAAAGGNATGTATTGGGNGGCGCCATTGAATACAGATGGGAGACTGTTG
GAGTATTATAGACTGTACAACCCACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTT
GCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAACA
TGTACAACACCGGNATATTGCCAGAGTTAACCTGACC

FIGURE 40

TCTCGCAGATAGTAAATAATCTCGGAAAGGCGAGAAAGAAGCTGTCTCCATCTTGTCTGTAT
CCGCTGCTCTTGTGACGTTGTGGAGATGGGGAGCGTCCTGGGGCTGTGCTCCATGGCGAGCT
GGATACCATGTTTGTGTGGAAGTGCCCCGTGTTTGCTATGCCGATGCTGTCTAGTGGAAC
AACTCCACTGTAACTAGATTGATCTATGCACTTTTCTTGCTTGTGGAGTATGTGTAGCTTG
TGTAATGTTGATACCAGGAATGGAAGAACAACCTGAATAAGATTCTGGATTTTGTGAGAAATG
AGAAAGGTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTT
GGTTTGGCTATGTTCTATCTTCTTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGA
TCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTA
TTATTGGGGCATTCTTCATTCCAGAAGGAACTTTTACAACCTGTGTGGTTTTATGTAGGCATG
GCAGGTGCCCTTTTGTTCATCTCATCAACTAGTCTTACTTATTGATTTTGCACATTCATG
GAATGAATCGTGGGTTGAAAAAATGGAAGAAGGGAACCTCGAGATGTTGGTATGCAGCCTTGT
TATCAGCTACAGCTCTGAATTATCTGCTGTCTTTAGTTGCTATCGTCCTGTTCTTTGTCTAC
TACACTCATCCAGCCAGTTGTTTCAAGAAACAAGGCGTTTATCAGTGTCAACATGCTCCTCTG
CGTTGGTGCTTCTGTAATGTCTATACTGCCAAAAATCCAAGAATCACAACCAAGATCTGGTT
TGTTACAGTCTTCAGTAATTACAGTCTACACAATGTATTTGACATGGTCAGCTATGACCAAT
GAACCAGAAACAAATTGCAACCCAAGTCTACTAAGCATAATTGGCTACAATACAACAAGCAC
TGTCCCAAAGGAAGGGCAGTCAGTCCAGTGGTGGCATGCTCAAGGAATTATAGGACTAATTC
TCTTTTTGTGTGTGTATTTTATTCCAGCATCCGTACTTCAAACAATAGTCAGGTTAATAAA
CTGACTCTAACAAGTGATGAATCTACATTAATAGAAGATGGTGGAGCTAGAAGTGATGGATC
ACTGGAGGATGGGGACGATGTTTACCGAGCTGTAGATAATGAAAGGGATGGTGTCACTTACA
GTTATTCCTTCTTTCACTTCATGCTTTTCTGGCTTCACTTTATATCATGATGACCCTTACC
AACTGGTCCAGGTATGAACCTCTCGTGAGATGAAAGTCAGTGGACAGCTGTCTGGGTGAA
AATCTCTTCCAGTTGGATTGGCATCGTGCTGTATGTTTGGACACTCGTGGCACCCTTGTTC
TTACAAATCGTGATTTTGACTGAGTGAAGTCTTAGCATGAAAGTCCCACCTTTGATTATTGC
TTATTGAAAACAGTATTTCCCACTTTTGTAAAGTTGTGTATGTTTTTGTCTCCCATGTAA
TTCTCCAGTGTTCTGGCATGAATTAGATTTTACTGCTTGTCAATTTGTTTATTTCTTACCAA
GTGCATTGATATGTGAAGTAGAATGAATTGCAGAGGAAAGTTTTATGAATATGGTGATGAGT
TAGTAAAGTGCGCATTATTGGGCTTATTCTCTGCTCTATAGTTGTGAAATGAAGAGTAAAA
ACAAATTTGTTTGACTATTTTAAATTTATATTAGACCTTAAGCTGTTTTAGCAAGCATTA
GCAAATGTATGGCTGCCCTTTTGAATATTTGATGTGTTGCCCTGGCAGGATACTGCAAGA
ATGGTTTTATTTTAAATTTTATAACAAGTCACTTAAATGCCAGTTGTCTGAAAAATCTTATA
AGGTTTTACCCTTGATACGGAATTTACACAGGTAGGGAGTGTTAGTGGACAATAGTGTAGGTTA
TGGATGGAGGTGTCCGTACTAAATTGAATAACGAGTAAATAATCTTACTTGGGTAGAGATGG
CCTTTGCCAACAAAGTGAAGTGTTTTGGTTGTTTTAACTCATGAAGTATGGGTTTCAAGTGA
AATGTTTGGAACTCTGAAGGATTTAGACAAGGTTTTGAAAAGGATAATCATGGGTTAGAAGG
AAGTGTTTTGAAGTCACTTTGAAAGTTAGTTTTTGGGCCAGCACGGTAGCTCACCCTTGGT
AATCCCAGCACTTTGGGAGCTTAAGTGGGTAGATTACTTGAGCCAGGAATTCAGACCAGCT
TGGCACATGGTGAACCTGTTCTATAAAAAATAATCTGGCTTTGAGCATATGCCTGTGGTCCAG
CACTGAGAGGCTAGTGAAGATTGCTGAGCCAGAGCCAAAGGTTGCAGTGAGCAAGTCACGT
CACTGCACTCTAGCTGGCACAGAGTAAGCCAAAAAATATATATATATTGAAATCAAGGAGG
CAAAATTTTGACAGGGAAGGAAGTAAGTCAAAACCCTAGGCTTTAGTAGGTACTTATATA
AAATCTAGTCCAGTTCTCTCATTTAAAAAATGAAGACACTGAAATACAGACTTAAATAGCT
CAGATAGCTAATTAGGAAATTTCAAGTTGGCCAATAATAGCATTCTCTCTGACATTTAAAA
TAATTTCTATTCAAAATACATGCATATTGATTTACACCTCATACTGTGATAATTAATGTGAT
GTGGATTGCTGGTGTCCAGCATGACCATAAACAGGTGAGAAGAATGATGGAATGTTTTAGA
ATAAACTCCTGCTTATAGTATACTACACAGTTCAAAGATGTTTTAAATGCTTTTGTATTTA
CTGCCATGTAATTGAAATATATAGATTATTGTAACCTTTCAACCTGAAATCAAGCAGTATG
AGAGTTTAGTTATTTGTATGTGTCACTAGTGTCTAATGAAGCTTTTAAATCTACAATTTCT
TCTTTAAAAATATTTATTAATGTGAATGGAATATAACAATTCAGCTTAATTTCCCAACCTTA
TTCTGTGTGTAGACATTGTATTCCACAATTTTGAATGGCTGTGTTTTACCTCTAAATAAATG
AATTCAGAGAAAAA

42/3-10

FIGURE 41

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGME
EQLNKIPGFCENEKGVVPCNILVG YKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNG
FWFFKFAAAIAIIIGAFFIPEGTFTTVWFYVGMAGAFCFILIQLVLLIDFAHSWNESWVEKM
EEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPASCS ENKAFISVNMLLCVGASVMSI
LPKIQESQPRSGLLQSSVITVYTM YLTWSAMTNEPETNCNPSLLSIIGYNTTSTVPKEGQSV
QWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLI EDGGARSDGSLEDGDDVH
RAVDNERDGV TYSYSFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVVWKISSSWIGI
VLT VWTLVAPLVLTNRDFD

43/310

FIGURE 42

GCGAGAAAGAAGCTGTCTCCATCTTGTCTGTATCCCGCTGCTTCTTGNGACGTTGTGGAGAT
GGGGAGCGTCCCTGGGGCTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCC
CCGTGTTTGCTATGCCGATGCTGTCTAGTGGAACAANTCCACTGTAACTAGATTGATCTA
TGCACTTTTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATAACCAGGAATGGAAG
AACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATT
TTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCT
CTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGAT
TTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

44/3-10

FIGURE 43

GTTATTGTGAACTTTGTGGAGATGGGAGGTCNTGGGGCTGTGTTCCATGGCGAGCTGGATAC
CANGTTTGTGTGGAAGTGCCCCGTGTTTGNTATGCCGATGCTGTCCTAGTGGAAACAANTCC
ACTGTAATTAGATTGATNTATGCACTTTTNTTGCTTGTTGGAGTANGTG TAGCTTGTGTAAT
GTTGATACCAGGAATGGAAGAACA ACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAG
GTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATNGTTTGTGCTTTGGTTTG
GCTANGTTCTATNTTCTTCTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAG
AGCTGCAGTGCACAATGGATTTTGGTTTTTTAAATTGCTGCAGCAATTGCAATTATTATTG
GGGC

FIGURE 44

AAGAAGCTGTCTCCATCTTGTCTGTATCCGCTGCTCTTGTGAACGTTNTGGAGATGGGGAGC
GTCCTTGGGGTTGTGCTCCATGGCGAGCTGGATAACCATGTTTGTGTGGAAGTGCCCCGTGTT
TGCTATGCCGATGCTGTCCTAGTGGAAACAACCTCCACTGTAAC TAGATTGATCTATGCACTT
TTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAGAACAAC
GAATAAGATTCCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATTTTGGTTG
GCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTTTA
CTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTT
CTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

FIGURE 45

GCTGTCCTTAGTGGAAACAANTCCAACCTTGTAACCTGGATTGATCTATGCACTTTTTTCCTTG
CTTGTTGGAGTATGTGTAGCTTTGTGTAATGTTGTTCCCAGGATTGGANGAACAACTGAATA
AGATTCCTGGATTTTTTGTGAGAATGAGAAAGGTGTTGTCCCCTTGTAACATTTTTTGGTTGGC
TATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTCTTTACT
AATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCT
TTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGCATTCTTCATTCCAGAAGGAACTTTT
ACAACTGTGTGGTTTTATGTAGGCATGGCAGGTGCCTTTTGTTCATCCTCATACAACTAGT
CTTACTTATTGATTTTGCACATTCATGGAATGAATCGTGGGTTGAAAAAATGGAAGAAGGGA
ACTCGAGATGTTGGTATGCAGCCTTGTTATCAGCTACAGCTCTGAATTATCTGCTGTCTTTA
GTTGCTATCGTCCTGTTCTTTGTCTACTACACTCATCCAGCCAGTTGTTTCAGAAAACAAGGC
GTTTCATCAGTGTCAACATGCTCCTCTGCGTTGGTGCTTCTGTAATG

FIGURE 46A

CTCGGGCGCGCACAGGCAGCTCGGTTTGCCCTGCGATTGAGCTGCGGGTCGCGGCCGGCGCC
GGCCTCTCCAATGGCAAATGTGTGTGGCTGGAGGCGAGCGCGAGGCTTTCGGCAAAGGCAGT
CGAGTGTTCAGACCGGGCGAGTCCTGTGAAAGCAGATAAAAGAAAACATTTATTAACGT
GTCATTACGAGGGGAGCGCCCGGGGGGCTGTGCGACTCCCCGCGGAACATTTGGCTCCCT
CCAGCTCCGAGAGAGGAGAAGAAGAAAGCGGAAAAGAGGCAGATTACGTCGTTTCCAGCCA
AGTGGACCTGATCGATGGCCCTCCTGAATTTATCACGATATTTGATTTATTAGCGATGCCCC
CTGGTTTGTGTGTTACGCACACACACGTGCACACAAGGCTCTGGCTCGCTTCCCTCCCTCGT
TTCCAGCTCCTGGGCGAATCCCACATCTGTTTCAACTCTCCGCCGAGGGCGAGCAGGAGCGA
GAGTGTGTCGAATCTGCGAGTGAAGAGGGACGAGGGAAAAGAAACAAAGCCACAGACGCAAC
TTGAGACTCCCGCATCCCAAAGAAGCACCAGATCAGCAAAAAAGAAGATGGGCCCCCGA
GCCTCGTGTGTGCTTGCTGTCCGCAACTGTGTTCTCCCTGCTGGGTGGAAGCTCGGCCTTC
CTGTGCGACCACCGCCTGAAAGGCAGGTTTCAGAGGGACCGCAGGAACATCCGCCCCAACAT
CATCCTGGTGTGCTGACGGACGACCAGGATGTGGAGCTGGGTTCATGCAGGTGATGAACAAGA
CCCCGCGCATCATGGAGCAGGGCGGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATG
TGCTGCCCCCTCACGCTCCTCCATCCTCACTGGCAAGTACGTCCACAACCACAACACCTACAC
CAACAATGAGAACTGCTCCTCGCCCTCCTGGCAGGCACAGCACGAGAGCCGCACCTTTGCCG
TGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAGTATCTTAATGAATACAAC
GGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTTCGACTCCTTAAAAACTCCCGCTTTTA
TAACTACACGCTGTGTGCGAACGGGGTGAAAGAGAAGCACGGCTCCGACTACTCCAAGGATT
ACCTCACAGACCTCATACCAATGACAGCGTGAGCTTCTTCCGCACGTCCAAGAAGATGTAC
CCGCACAGGCCAGTCCTCATGGTCATCAGCCATGCAGCCCCCAGGCCCTGAGGATTCAGC
CCCACAATATTACGCCTCTTCCCAAACGCATCTCAGCACATCAGCCGAGCTACAACTACG
CGCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGCCCATGAAGCCCATCCACATG
GAATTCACCAACATGCTCCAGCGGAAGCGCTTGCAGACCCTCATGTGCGGTGGACGACTCCAT
GGAGACGATTTACAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTACATCGTATACA
CCGCCGACCACGGTTACACATCGGCCAGTTTGGCCTGGTGAAAGGGAAATCCATGCCATAT
GAGTTTGACATCAGGGTCCCGTTCTACGTGAGGGGCCCCAACGTGGAAGCCGGCTGTCTGAA
TCCCCACATCGTCCTCAACATTGACCTGGCCCCCACCATCCTGGACATTGCAGGCCTGGACA
TACCTGCGGATATGGACGGGAAATCCATCCTCAAGCTGCTGGACACGGAGCGGCCGGTGAAT
CGGTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGGAGAGAGGCAA
GCTGCTACACAAGAGAGACAATGACAAGGTGGACGCCAGGAGGAGAACTTTCTGCCCAAGT
ACCAGCGTGTGAAGGACCTGTGTGAGCGTGTGAGTACCAGACGGCGTGTGAGCAGCTGGGA
CAGAAGTGGCAGTGTGTGGAGGACGCCACGGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCC
CATGCGGCTGGGCGGCAGCAGAGCCCTCTCAAACCTCGTGCCCAAGTACTACGGGCAGGGCA
GCGAGGCCTGCACCTGTGACAGCGGGGACTACAAGCTCAGCCTGGCCGGACGCCGGAAAAAA
CTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCAGTCGCTCCATCCGCTCAGTGGCCAT
CGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCAGCCCCGAAACCTCA
CCAAGCGGCACTGGCCAGGGGCCCCCTGAGGACCAAGATGACAAGGATGGTGGGGACTTCAGT
GGCACTGGAGGCCTTCCCGACTACTCAGCCGCCAACCCATTAAAGTGACACATCGGTGCTA
CATCCTAGAGAACGACACAGTCCAGTGTGACCTGGACCTGTACAAGTCCCTGCAGGCCTGGA
AAGACCACAAGCTGCACATCGACCACGAGATTGAAACCCTGCAGAACAAAATTAAGAACCTG
AGGGAAGTCCGAGGTACCTGAAGAAAAAGCGGCCAGAAGAATGTGACTGTCAAAAATCAG
CTACCACACCCAGCACAAAGGCCGCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTCAGGA
AGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAGAACTC
CGCAAGCTGCTCAAGCGCCTGCAGAACAACGACACGTGCAGCATGCCAGGCCTCACGTGCTT
CACCCACGACAACCAGCACTGGCAGACGGCGCCTTCTGGACACTGGGGCCTTTCTGTGCCT
GCACCAGCGCCAACAATAACACGTACTGGTGCATGAGGACCATCAATGAGACTCACAAATTC

FIGURE 46B

CTCTTCTGTGAATTTGCAACTGGCTTCCTAGAGTACTTTGATCTCAACACAGACCCCTACCA
GCTGATGAATGCAGTGAACACACTGGACAGGGATGTCCTCAACCAGCTACACGTACAGCTCA
TGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAACCCCGGACTCGAAACATGGACCTG
GATGGAGGAAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAG
ACCTTCTTCCAAATCACTGGGACAACTGTGGGAAGGCTGGGAAGGTTAAAGAAACAACAGAGG
TGGACCTCCAAAAACATAGAGGCATCACCTGACTGCACAGGCAATGAAAAACCATGTGGGTG
ATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCCTGAGAAAGCAAGCACGCACTCTCAGTC
AACATGACAGATTCTGGAGGATAACCAGCAGGAGCAGAGATAACTTCAGGAAGTCCATTTTT
GCCCCTGCTTTTGCTTTGGATTATACCTCACCAGCTGCACAAAATGCATTTTTTTCGTATCAA
AAAGTCACCACTAACCCCTCCCCCAGAAGCTCACAAAGGAAAACGGAGAGAGCGAGCGAGAGA
GATTTCTTGAAATTTCTCCCAAGGGCGAAAGTCATTGGAATTTTAAATCATAGGGGAAA
AGCAGTCCTGTTCTAAATCCTCTTATTCTTTTGGTTTGTCAAAAGAAGGAAGTAAGAAGCA
GGACAGAGGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACGTTTGACAATGAGTCAGTAGC
ACAAAAGAGATGACATTTACCTAGCACTATAAACCCCTGGTTGCCTCTGAAGAACTGCCTTC
ATTGTATATATGTGACTATTTACATGTAATCAACATGGGAACCTTTTAGGGGAACCTAATAAG
AAATCCCAATTTTCAGGAGTGGTGGTGTCAATAAACGCTCTGTGGCCAGTGTAAAAGAAAAA

FIGURE 47

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQDRDRNIRPNIILVLTDDQDVELGSMQ
VMNKTRRIMEQGGAHFINAFVTTPMCCPSRSSILTGKYVHNHNTYTNNECSPSWQAQHE
RTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTLCRNGVKEKHGSD
YSKDYLTDLITNDSVSFFRTSKKMPHRPVLMMVISHAAPHGPEDSAPQYSRLFPNASQHITP
SYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQQRKRLQTLMSVDDSMETIYNMLVETGELDNT
YIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDI
AGLDIPADM DGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEEN
FLPKYQRVKDLCQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKY
YGQGSEACTCDSDYKLSLAGRRKKLFKKKYKASYVRSRSIRSV AIEVDGRVYHVGLGDAAQ
PRNLTKRHWPGAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYILENDTVQCDLDLYKS
LQAWKDHLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGR LKHGSSSL
HPFRKGLQEKDKVWLLREQRKKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLG
PFCACTSANNNTYWCMTINETHNFLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQL
HVQLMELRSCKGYKQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

50/310

FIGURE 48

AACAAAGTTCAGTGAAGGCTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGA
GCTGCTGGGCAGAGAGGGACTGTCCGGCTCCCAGATGCTGGGCCTCCTGGGGAGCACAGCCC
TCGTGGGATGGATCACAGGTGCTGCTGTGGCGGTCTGCTGCTGCTGCTGCTGCTGGCCACC
TGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGGAACCGTACAGCTGCAGGGGGAAA
CCGAGTCCGCCGGGCCAGCCTTGGCCCTTCCGGCGGGGGCCACCTGGGAATCTTTCACC
ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC
CCCCGCCACACCCCTCACCACCTCCACCACCACCACCACCCCCACCGCCACCATCCCCGCCA
CGCTCGCTGAGGCTGCTGTGCGCGGTGCCTGTGGACAGCAGCTGCCCCCTGCCCTCCCATCTG
TTCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG
G3AACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTACTATGGC
AGCAGAGAATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGGAGAGCAAGG
GTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGAAGTCCCAGTGAGCCCCAGAAATG
ACAAGCGTGTCTTGGCAGAGCCAGCACACAAGTGGATGTGAAGTGCCCGTCTTGACCTCCTC
ATCAGGCTGCTGCAGGCCTCTGGCGGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCTTTT
GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTCATCGGTGCCTTAGTCCAAGAAAAT
AAAAACCACTAAGAAGCTTTAAAAAAAAAAAAAAAAAAAAA

FIGURE 49

MLGLLGSTALVGWITGAAVAVLLLLLLLATCLFHGRQDCDVERNRTAAGGNRVRRAPWPFR
RRGHLGIFHHHRHPGHVSHVPNVGLHHHHHPRHTPHHLHHHHHPRHHPRHAR

52/310

FIGURE 50

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTGGGGATCCAGAGCCATGTCGGACCTGCTA
CTACTGGGCCTGATTGGGGCCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGG
GTACTCAGGGCTACTGGCTGGGGTGGAAGTGAGTGCTGGGTACCCCCCATCCGCAACGTCA
CTGTGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGC
TGCAGCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCC
CCCTGATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCC
CTGAGCTCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGC
CATGTGGTGACAGCCACCTTCCCCTACACCACCATTCTGTCCATCTGGCTGGCTACCCGCCG
TGTCCATCCTGCCTTGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGG
AGATCTACCAGGAAGACCAGATCCATTTTATGTGCCCACTGGCACGGCAGGGAGACTTCTAT
GTGCCTGAGATGAAGGAGACAGAGTGGAATGGCGGGGGCTTGTGGAGGCCATTGACACCCA
GGTGGATGGCACAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGAAGTGAGCC
CTGGCAGCCGGGAGACTTCAGCTGCCACACTGTACCTGGGGCGAGCAGCCGTGGCTGGGAT
GACGGTGACACCCGAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGA
GGAGCTGGACTTGGAGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGC
CCCTGGGGACTACCAAGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCAAGGAGTAACCC
ATGGCCTGCACCCTCCTGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCT
CCAGCCCTCTTCCTCCTTCTCTGGGGGAGGAGGGGTTCTGAGGGACCTGACTTCCCCTGC
TCCAGGCCTCTTGCTAAGCCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCA
GGGACTATTTTCTGCACCAGCCCCCAGGGCTGCCGCCCTGTTGTGTCTTTTTTTCAGACTC
ACAGTGGAGCTTCCAGGACCCAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAA
AAAAA

53/310

FIGURE 51

MSDLLLLGLIGGLTLLLLLTLAFAAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGR
LFTESCSISPKLRSIAVYYDNPHMVPPDKCRAVGSILSEGEESPSPELIDLYQKFGFKVFS
FPAPSHVVTATFPYTTILSIWLATTRVHPALDTYIKERKLCAYPRLEIYQEDQIHMCPLAR
QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAA TLSPGAS
SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGLGESRLDPGTEPLGTTKWLWEPTAPEKGKE

54/310
FIGURE 52

CCGCGGGAACGCTGTCCTGGCTGCCGCCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCT
GCCCCGCGCCAGTCATGACCCTGCGCCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCT
GCTGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGA
CCCTCCAAGTGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGA
GACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCT
GACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGA
GTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCTTCTCACTTGGCCTAT
GGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCT
GATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGGGCATTGCTGCTGCTGCTAG
GGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAAT
AGACCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAAATA
ATAAATAATAAATTTAAAAAACTTAAAAAAAAAAAAAAAAAAAA

55/310

FIGURE 53

MTLRPSLLPLHLLLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFCDTLHI
HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLDMCVGEKRRAIIPSHLAYGKRGF
PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS
KKKLKEEKRNKSKKK

56/310

FIGURE 54

CCCGGGAACGTGTTCCCTGGCTGCCGCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCTGC
CCCGCGCCCAGTCATGACCCTGCGCCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCTGC
TGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGACC
CTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGAGA
CACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCTGA
CCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGAGT
CTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCTTCTCACTTGGCCTATGG
AAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAAGTATGACGTGGAGCTGA
TTGCACTAATCCGAGCCAACACTACTGGCTAAAGCTGGTGAAGGGCATTTTGCCTCTGGTAGGG
ATGGCCATGGTGCCACCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAATAGA
CCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAAATAATA
AATAATAAATTTTAAAAAACTTA

57/310

FIGURE 55

CCGAAAGTCCCGTCCGGACCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACG
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGA
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATT
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCACCATCTGTCCCAGCGGATGCAGTGGT
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGG
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAA
CAAGAGCAAAAAGAAATAATAAATAATAAATTTTAAAAAACTTAAAA

58/310

FIGURE 56

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTTTTCTTGGCTAAAATCGGGGGAG
TGAGGCGGGCCGGCGCGGCGGACACCGGGCTCCGGAACCACTGCACGACGGGGCTGGACTG
ACCTGAAAAAAATGTCTGGATTTCTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGG
GAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGAT
TATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCT
GTGGTGTATATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGA
GGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGG
TTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTG
CTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATCTTT
TTTGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGACTTATGGCAGTGAACACATCTGAT
TTCCACAGCACAACAGCCCTGCATGGGTTTGTGTTTTTTTACTGCTCACTCCCAACCTT
TTGTAATGCCATTTTCTAAACTTATTTCTGAGTGTAGTCTCAGCTTAAAGTTGTGTAATACT
AAAATCACGAGAACACCTAAACAACAACCAAAAATCTATTGTGGTATGCACTTGATTAACCT
ATAAAATGTTAGAGGAACTTTACATGAATAATTTTTGTCAAATTTTATCATGGTATAATT
TGTA AAAATAAAAAGAAATTACAAAAGAAATTATGGATTTGTCAATGTAAGTATTTGTCATA
TCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGTGTTTATTCAAATGTGGT
CTCTTCTGTGTCAAATGTTAAATGAAATATAAACATTTTTTAGTTTTTAAATATTCCGTGG
TCAAATTCCTTCCCTCACTATAATTGGTATTTACTTTTACCAAAAATTCTGTGAACATGTAAT
GTA ACTGGCTTTTGAGGGTCTCCCAAGGGTGAGTGGACGTGTTGGAAGAGAGAAGCACCAT
GGTCCAGCCACCAGGCTCCCTGTGTCCCTTCCATGGGAAGGTCTTCCGCTGTGCCTCTCATT
CCAAGGGCAGGAAGATGTGACTCAGCCATGACACGTGGTTCTGGTGGGATGCACAGTCACTC
CACATCCACCACTG

59/310
FIGURE 57

MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWIIIDA AVIYPTMKDFNHSYHACGVI
ATIAFLMINAVSNGQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEK
DIVYPGIAVFFQNAFIFFGGLVFKFGRTEDLWQ

60/310
FIGURE 58

TTCTTGGCTAAAATCGGGGGAGTGAGGCGGGCCGGCGCGGCGGACACCGGGCTCCGGAACC
ACTGCACGACGGGGCTGGACTGACCTGAAAAAATGTCTGGATTTCTAGAGGGCTTGAGATG
CTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTAC
TATTTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGAT
TTCAACCACTCATACCATGCCTGTGGTGTATAGCAACCATAGCCTTCCTAATGATTAATGC
AGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTG
CTCGCATTGCGCTTTTCGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGG
ATTCTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATT
TTTCCAGAATGCCTTCATCTTTTTTTGGAGGGCTGGTTTTTAAGTTTGGC

61/310
FIGURE 59

TGGACGGACCTGAAAAAATGTTTGGATTTNTAGAGGGNTTGAGATGTTTCAGAATGCATGAC
TGGGGGAAAAGCGCAAATACTATTGCTTCCATTGCTGCTGGTGTANTATTTTTTACAGGCTG
GTGGATTATCATAGATGCAGNTGTTATTTATCCCACCATGAAAGATTTCAACCANTCATACC
ATGCCTGTGGTGTATAGCAACCATAGCCTTCNTAATGATTAATGCAGTATCGAATGGACAA
GTCCGAGGTGATAGTTACAGTGAAGGTTGTTTGGGTCAAACAGGTGCTCGCATTTGGCTTTT
CGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTT
ATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGNTGTATTTTTCCAGAATGCCTTC
ATCTTTTTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGANTTATGGCAGTG

62/310

FIGURE 60

GGACACCGGGTTCCGGACCAATGCANGACGGGGTGGANTGACCTGAAAAAATGTTTGGATT
TTTAGAGGGCTTGAGATGNTCAGAATGCATTGACTGGGGGAAAAGCGCAATANTATTGCTTT
CCATTGCTGCTGGTGTACTATTTTTTACAGGGTGGTGGATTATCATAGATGCAGCTGTTATT
TATCCCACCATGAAAGATTTNAACCACTCATACCATGCCTGTGGTGTTATAGCAACCATAGC
CTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTT
GTTTGGGTCAAACAGGTGNTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGGATTT
CTGATTGNATTCTATGCGGATTCTTCTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTAT
ACCCTGGAATTNCTNTATTTTTTCCAGAATGCC

63/310

FIGURE 61

TAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATANTATTGCTTCC
ATTGNTGNTGGTGTANTATTTTTTTTACAGGCTGGTGGATTATNATAGATGCAGCTGTTATTT
ATCCCACCATGAAAGATTTNAACCANTCATACCATGCCTGTGGTGTTATAGCAACCATAGCC
TTCCTAATGATTAATGCAGTATNGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTG
TTTGGGTCAAACAGGTGNTNGCATTTGGCTTTTNGTTGGTTTCATGTTGGCCTTTGGATCTN
TGATTGCATTTATGTGGATTNTTTTTTGGAGGTTATGTTGCTAAAGNAAAAGACATAGTATAC
CCTGT

64/310

FIGURE 62

GGGAGGCTGTGNCCGTTTTGTTTTNTTGGCTAAAATCGGGGGAGTGAGGCGGCCCGGCGCGG
CGNGACACCGGGTTCCGGGAACCATTCACGACGGGGTGGACTGACCTGAAAAAATGTTTG
GATTNTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATT
GCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGT
TATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCTGTGGTGTATAGCAACCA
TAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAA
GGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGG
ATNTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAG
TATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATNTTTTTTGGAGGGCTG

FIGURE 63

CGACGCCCGCGTGATGTGGCTTCCGCTGGTGCTGCTCCTGGCTGTGCTGCTGCTGGCCGTCC
TCTGCAAAGTTTACTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTC
AAACGGCCCCCAGCGCCCCCTGGTAACTGACAAGGAGGCCAGGAAGAAGGTTCTCAAACAAGC
TTTTTCAGCCAACCAAGTGCCGGAGAAGCTGGATGTGGTGGTAATTGGCAGTGGCTTTGGGG
GCCTGGCTGCAGCTGCAATTCTAGCTAAAGCTGGCAAGCGAGTCCTGGTGCTGGAACAACAT
ACCAAGGCAGGGGGCTGCTGTCATACCTTTGGAAAGAATGGCCTTGAATTTGACACAGGAAT
CCATTACATTGGGCGTATGGAAGAGGGGCAGCATTGGCCGTTTTATCTTGACCAGATCACTG
AAGGGCAGCTGGACTGGGCTCCCCTGTCCTCTCCTTTTGACATCATGGTACTGGAAGGGCCC
AATGGCCGAAAGGAGTACCCCATGTACAGTGGAGAGAAAGCCTACATTCAGGGCCTCAAGGA
GAAGTTTCCACAGGAGGAAGCTATCATTGACAAGTATATAAAGCTGGTTAAGGTGGTATCCA
GTGGAGCCCCCTCATGCCATCCTGTTGAAATTCCTCCCATTGCCCGTGGTTTCACTCCTCGAC
AGGTGTGGGCTGCTGACTCGTTTCTCTCCATTCTTCAAGCATCCACCCAGAGCCTGGCTGA
GGTCCTGCAGCAGCTGGGGGCTCCTCTGAGCTCCAGGCAGTACTCAGCTACATCTTCCCCA
CTTACGGTGTCAACCCCAACCAAGTGCCTTTTCCATGCACGCCCTGCTGGTCAACCACTAC
ATGAAAGGAGGCTTTTATCCCCGAGGGGGTCCAGTGAATTTGCCTTCCACACCATCCCTGT
GATTGAGCGGGCTGGGGGCGCTGTCTCACAAAGGCCACTGTGCAGAGTGTGTTGCTGGACT
CAGCTGGGAAAGCCTGTGGTGTCACTGTGAAGAAGGGGCATGAGCTGGTGAACATCTATTGC
CCCATCGTGGTCTCCAACGCAGGACTGTTCAACACCTATGAACACCTACTGCCGGGGAACGC
CCGCTGCCTGCCAGGTGTGAAGCAGCAACTGGGGACGGTGCGGCCCGGCTTAGGCATGACCT
CTGTTTTCTATCTGCCTGCGAGGCACCAAGGAAGACCTGCATCTGCCGTCCACCAACTACTAT
GTTTACTATGACACGGACATGGACCAGGCGATGGAGCGCTACGTCTCCATGCCCAGGGAAGA
GGCTGCGGAACACATCCCTCTTCTCTTCTTCTGCTTTCCCATCAGCCAAAGATCCGACCTGGG
AGGACCGATTCCAGGCCGGTCCACCATGATCATGCTCATACCCACTGCCTACGAGTGGTTT
GAGGAGTGGCAGGCGGAGCTGAAGGGAAAGCGGGCAGTGAATGAGACCTTCAAAAATC
CTTTGTGGAAGCCTCTATGTCAGTGGTCTGAAACTGTTCCACAGCTGGAGGGGAAGGTGG
AGAGTGTGACTGCAGGATCCCCACTCACCACAGTTCTATCTGGCTGCTCCCCGAGGTGCC
TGCTACGGGGCTGACCATGACCTGGGGCGCTGCACCTTGTGTGATGGCCTCCTTGAGGGC
CCAGAGCCCCATCCCCAACCTCTATCTGACAGGCCAGGATATCTTACCTGTGGACTGGTCG
GGGCCCTGCAAGGTGCCCTGCTGTGCAGCAGCGCCATCCTGAAGCGGAACCTGTACTCAGAC
CTTAAGAATCTTGATTCTAGGATCCGGGCACAGAAGAAAAAGAATTAGTTCCATCAGGGAGG
AGTCAGAGGAATTTGCCCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTG
CATTAGTTCCTTGACGTATAAAGCACTAATTTGGTTCTGATGCCTGAAGAGAGGCCTAG
TTTAAATCACAATCCGAATCTGGGGCAATGGAATCACTGCTTCCAGCTGGGGCAGGTGAGA
TCTTTACGCCTTTTATAACATGCCATCCCTACTAATAGGATATTGACTTGGATAGCTTGATG
TCTCATGACGAGCGGCGCTCTGCATCCCTCACCCTGCCTCCTAACTCAGTGATCAAAGCGA
ATATTCCATCTGTGGATAGAACCCTGGCAGTGTTGTGCTCAGCTCAACCTGGTGGGTTCAAGT
TGTCCTGAGGCTTCTGCTCTCATTCAATTTAGTGCTACGCTGCACAGTTCTACACTGTCAAGG
GAAAAGGGAGACTAATGAGGCTTAACCTCAAAACCTGGGCGTGGTTTTGGTTGCCATTCCATA
GGTTTGGAGAGCTCTAGATCTCTTTTGTGCTGGGTTCAAGTGGCTCTTCAAGGGGACAGGAAAT
GCCTGTGTCTGGCCAGTGTGGTTCTGGAGCTTTGGGGTAACAGCAGGATCCATCAGTTAGTA
GGGTGCATGTGAGATGATCATATCCAATTCATATGGAAGTCCCCGGTCTGTCTTCTTATCA
TCGGGGTGGCAGCTGGTTCTCAATGTGCCAGCAGGGACTCAGTACCTGAGCCTCAATCAAGC
CTTATCCACCAAATACACAGGGAAGGGTGATGCAGGGAAGGGTGACATCAGGAGTCAGGGCA
TGGACTGGTAAGATGAATACTTTGCTGGGCTGAAGCAGGCTGCAGGGCATTCCAGCCAAGGG
CACAGCAGGGGACAGTGCAGGGAGGTGTGGGGTAAGGGAGGGAAGTCACATCAGAAAAGGGA
AAGCCACGGAATGTGTGTGAAGCCAGAAATGGCATTTCAGTTAATTAGCACATGTGAGGG
TTAGACAGGTAGGTGAATGCAAGCTCAAGGTTTGAAAAATGACTTTTCAGTTATGTCTTTG
GTATCAGACATACGAAAGTCTCTTTGTAGTTTCGTGTTAATGTAACATTAATAAATTTATTG
ATTCCATTGCTTTAAAAA

66/310

FIGURE 64

MWLPLVLLLLAVLLLLAVLCKVYLGLFSGSSPNPFSEDVKRPPAPLVTDKEARKKVLKQAFSAN
QVPEKLDVVVIGSGFGGLAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLEFDTGIHYIG
RMEEGSIGRFILDQITEGQLDWAPLSSPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQ
EEAIDKYIKLVKVVSSGAPHAILLKFLPLPVVQLLDRCGLLTRFSPFLQASTQSLAEVLQQ
LGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHYMKGGFYPRGGSSEIAFH TIPVIQRA
GGAVLT KATVQSVLLDSAGKACGVS VKKGHELVNIYCPIVVS NAGLFNTYEHLLPGNARCLP
GVKQQ LGTVRPGLGMTSVFI CLRGT KEDLHLPSTNYYVYYDTMDQAMERYVSM PREEAAEH
IPLLFFAFPSAKDPTWEDRFPGRSTMIMLIPTAYEWFEWQAELKGKRGSDYETFKNSFVEA
SMSVLKLFPPQLEGKVESVTAGSPLTNQFYLAAPRGACYGADHDLGRLHPCVMASLRAQSPI
PNLYLTGQDIFT CGLVGALQGALLCSSAILKRNL YSDLKNLDSRIRAQKKKN

GCAGCGGCGAGGCGGCGGTGGTGGCTGAGTCCGTGGTGCCAGAGGCCGAAGGCGACAGCTCTA
GGGGTTGGCACCGGCCCCGAGAGGAGGATGCGGGTCCGGATAGGGCTGACGCTGCTGCTGTG
TGCGGTGCTGCTGAGCTTGGCCTCGGCGTCCTCGGATGAAGAAGGCAGCCAGGATGAATCCT
TAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTAAAGGACCATACTACTGCAGGCAGA
GTAGTTGCTGGTCAAATATTTCTTGATTCAGAAGAATCTGAATTAGAATCCTCTATTCAAGA
AGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTCACAGAAGATATCAGCTTTCTAG
AGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTACGGAAACCAGCTTTG
ACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCTTTTCCTAGATAA
GGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTGCTACAACCT
ATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAAACTGAAGAAGAGGCTGCTAAGAGA
CGGCAGATGCAGGAAGCAGAAATGATGTATCAAACCTGGAATGAAAATCCTTAATGGAAGCAA
TAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTCTATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATC
CAGGCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGC
TCTTGGCTTTCTGTATGCCTCTGGACTTGGTGTTAATTCAAGTCAGGCAAAGGCTCTTGAT
ATTATACATTTGGAGCTCTTGGGGGCAATCTAATAGCCACATGGTTTTTGGTAAGTAGACTT
TAGTGGAAGGCTAATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACTTTT
TCAGCTTTTCATGATCCAGATTTGCTTGATTAAGACCAAATATTCAGTTGAACTTCCTTCAA
ATTCTTGTTAATGGATATAACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACA
ATTTTTCTTTAAATGATTAGTTTGGCTGATTGCCCCATAAAAGAGAGATCTGATAAATGGC
TCTTTTTAAATTTTCTCTGAGTTGGAATTGTCAGAATCATTTTTTTACATTAGATTATCATAA
TTTTTAAATTTTTCTTTAGTTTTTCAAATTTTGTAAATGGTGGCTATAGAAAAACAACAT
GAAATATTATACAATATTTTGAACAATGCCCTAAGAATTGTAAATTCATGGAGTTATTT
GTGCAGAATGACTCCAGAGAGCTCTACTTTCTGTTTTTTACTTTTCATGATTGGCTGTCTTC
CCATTTATTCTGGTCATTTATTGCTAGTGACACTGTGCCTGCTTCCAGTAGTCTCATTTTCC
CTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGGAAGATTAACTCATTTTTTAATAAA
ATTATGTCTAAGATTAAAAA
AAAAA

68/310
FIGURE 66

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLSKTTLTSDSVKDHTTAGRVVAGQIFLD
SEESELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHG
EPCHFPLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEEAAKRRQMGEAEMM
YQTGMKILNGSNKKSQKREAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFEK
LTEEGSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLIAHMLVLSRL

69/310

FIGURE 67

CTTCCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAGAACTTCTACTTGCCT
GCCTCCCTGCCTCTGGCCATGGCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCT
GTCAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGG
CTCAACTCTCCTGCACGCTCAGCCCCAGCACGTCACCATCAGGGACTACGGTGTGTCCTGG
TACCAGCAGCGGGCAGGCAGTGCCCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCA
CCACCGGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCCAATGCCT
GTGTCCTCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGC
TACGGCTTTAGTCCCTTAGGGGTGGGGTGTGAGATGGGTGCCTCCCCTCTGCCTCCCATTTCT
GCCCCTGACCTTGGGTCCCTTTTAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAAATGGG
TTAATAATATTCAACATGTCAACAAC

FIGURE 68

MACRCLSFLLMGTFLSVSQTVLAQLDALLVFPQVAQLSCTLSPQHVTIRDYGVSWYQQRAG
SAPRYLLYYRSEEDHHRPADIPDRFSAKDEAHNACVLTISPVPEDDADYYCSVGYGFSF

#1/310

FIGURE 69

GCCGCCCCGCCCCGAGACCGGGCCCGGGGGCGCGGGGCGGCGGGATGCGGCGCCCCGGGGCGG
CGATGACCGCGGAGCGCACGCGCGGGCCCGGGCCCTGACCCCGCCGCCCCGCTGAGCCC
CCCCCGGAGGTCCGGACAGGCCGAGATGACGCGCGAGCCCCCTGTTGCTGCTCCTGCTGCCGC
CGCTGCTGCTGGGGGCTTCCACCGGGCCGCGCCGCCCCGAGGCCCCCCAAAGATGGCGGAC
AAGGTGGTCCCACGGCAGGTGGCCCGGCTGGGCCGCACTGTGCGGCTGCAGTGCCAGTGGGA
GGGGGACCCGCGCCGCTGACCATGTGGACCAAGGATGGCCGCACCATCCACAGCGGCTGGA
GCCGCTTCCGCGTGCTGCCGCAGGGGCTGAAGGTGAAGCAGGTGGAGCGGGAGGATGCCGGC
GTGTACGTGTGCAAGGCCACCAACGGCTTCGCGAGCCTGAGCGTCAACTACACCCTCGTCGT
GCTGGATGACATTAGCCCAGGGAAGGAGAGCCTGGGGCCCCGACAGCTCCTCTGGGGGTCAAG
AGGACCCCGCCAGCGCAGTGGGCACGACCGCGCTTCACACAGCCCTCCAAGATGAGGCGC
CGGGTGATCGCACGGCCCGTGGGTAGCTCCGTGCGGCTCAAGTGCGTGCCAGCGGGCACCC
TCGGCCCCGACATCACGTGGATGAAGGACGACCAGGCCTTGACGCGCCAGAGGCCGCTGAGC
CCAGGAAGAAGAAGTGGACACTGAGCCTGAAGAACCTGCGGCCGAGGACAGCGGCAATAC
ACCTGCCGCGTGTGAAACCGCGCGGGCGCCATCAACGCCACCTACAAGGTGGATGTGATCCA
GCGGACCCGTTCCAAGCCCGTGCTCACAGGCACGCACCCCGTGAACACGACGGTGGACTTCG
GGGGGACCACGTCTTCCAGTGCAAGGTGCGCAGCGACGTGAAGCCGGTGATCCAGTGGCTG
AAGCGCGTGGAGTACGGCGCCGAGGGCCGCCACAACCTCCACCATCGATGTGGGCGGCCAGAA
GTTTGTGGTGCTGCCCCAGGGTGACGTGTGGTTCGCGGCCCGACGGCTCCTACCTCAATAAGC
TGCTCATACCCGTGCCCGCCAGGACGATGCGGGCATGTACATCTGCCTTGCGCCAAACACC
ATCGGCTACAGCTTCCGCGAGCGCTTCCCTACCGTGCTGCCAGACCCAAAACGCCAGGGCC
ACCTGTGGCCTCCTCGTCTCGGCCACTAGCCTGCCGTGGCCCGTGCTCATCGGCATCCAG
CCGGCGCTGTCTTCATCTGGGCACCCCTGCTCCTGTGGCTTTGCCAGGCCAGAAGAAGCCG
TGCACCCCCGCGCCTGCCCTCCCTGCTGGGCACCGCCCGCCGGGGACGGCCCGCGACCG
CAGCGGAGACAAGGACCTTCCCTCGTTGGCCGCCCTCAGCGCTGGCCCTGGTGTGGGGCTGT
GTGAGGAGCATGGGTCTCCGGCAGCCCCCAGCACTTACTGGGCCAGGCCAGTTGCTGGC
CCTAAGTTGTACCCCAAACCTCTACACAGACATCCACACACACACACACACACTCTCACAC
ACACTCACACGTGGAGGGCAAGGTCCACCAGCACATCCACTATCAGTGCTAGACGGCACCGT
ATCTGCAGTGGGCACGGGGGGGCGGCCAGACAGGCAGACTGGGAGGATGGAGGACGGAGCT
GCAGACGAAGGCAGGGGACCCATGGCGAGGAGGAATGGCCAGCACCCAGGCAGTCTGTGTG
TGAGGCATAGCCCCCTGGACACACACACACAGACACACACACTACCTGGATGCATGTATGCAC
ACACATGCGCGCACACGTGCTCCCTGAAGGCACACGTACGCACACGCACATGCACAGATATG
CCGCTGGGCACACAGATAAGCTGCCCAAATGCACGCACACGCACAGAGACATGCCAGAACA
TACAAGGACATGCTGCCTGAACATACACACGCACACCCATGCGCAGATGTGCTGCCTGGACA
CACACACACACACGGATATGCTGTCTGGACGCACACACGTGCAGATATGGTATCCGGACA
CACGTGCACAGATATGCTGCCTGGACACACAGATAATGCTGCCTTGACACACACATGCACGG
ATATTGCCTGGACACACACACACACACACACCGGTGCACAGATATGCTGTCTGGACACGCAC
ACATGCAGATATGCTGCCTGGACACACACTTCCAGACACACGTGCACAGGGCGCAGATATGCT
GCCTGGACACACGCAGATATGCTGTCTAGTCACACACACACGCAGACATGCTGTCCGGACAC
ACACACGCATGCACAGATATGCTGTCCGGACACACACACGCACGCAGATATGCTGCCTGGAC
ACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGATATTGCCTGGACACACACA
TGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATATGCTGTCCGGATACACACG
CACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGACACACATGCACACACAGGT
GCAGATATGCTGCCTGGACACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGA
TATTGCCTGGACACACACATGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATA
TGCTGTCCGGATACACACGCACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGA
CACACATGCACACACAGGTGCAGATATGCTGCCTGGACACACGCAGACTGACGTGCTTTTGG
GAGGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCATAGTTGATGAGGGACTTT
CCCTGCTCCACCGTCACTCCCCCAACTCTGCCCGCCTCTGTCCCCGCTCAGTCCCCGCTC
CATCCCCGCTCTGTCCCTGGCCTTGGCGGCTATTTTTGCCACCTGCCTGGGTGCCCGAGG
AGTCCCCTACTGCTGTGGGCTGGGGTTGGGGGCACAGCAGCCCCAAGCCTGAGAGGCTGGAG
CCCATGGCTAGTGGCTCATCCCCAGTGCACTTCTCCCCCTGACACAGAGAAGGGGCCTTGGTA
TTTATATTTAAGAAATGAAGATAATATTAATAATGATGGAAGGAAGACTGGGTTCAGGGAC
TGTGGTCTCTCCTGGGGCCCCGGGACCCGCTGGTCTTTCAGCCATGCTGATGACCACACCCC
GTCCAGGCCAGACACCACCCCCCACCCCACTGTCTGTGGTGGCCCCAGATCTCTGTAATTTTA
TGTAAGTTTTGAGCTGAAGCCCCGTATATTTAATTTATTTTGTAAACACAAAA

72/310

FIGURE 70

MTPSPLLLLLLPPLLLGAFPPAAAARGPPKMADKVVPQRVARLGRTVRLQCPVEGDPPLTM
WTKDGRTIHSGWSRFRVLPQGLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVVLLDDISPGK
ESLGPDSSSSGGQEDPASQQWARPRFTQPSKMRRRVVIARPVGSSVRLKCVASGHPRPDITWMK
DDQALTRPEAAEPRKKKWTLSLKNLRPEDSGKYTCRVSNRAGAINATYKVDVIQRTRSKPVL
TGTHPVNTTVDFGGTTSFQCKVRSADV KPV IQWLKRVEYGAEGRHNSTIDVGGQKFVVLPTGD
VWSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRSAFLTVLPDPKPPGPPVASSSSA
TSLPWPVVIGIPAGAVFILGTLLLWLCQAQKKPCTPAPAPPLPGHRPPGTARDRSGDKDLPS
LAALSAGPGVGLCEEHGSPAAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTSHHTSHSHVEGKV
HQHIHYQC

73/310

FIGURE 71A

CCCAGCTGAGGAGCCCTGCTCAAGACACGGTCACTGGATCTGAGAACTTCCCAGGGGACCG
CATTCCAGAGTCAGTGAAGCACCACATCTACCTCTTGCCACGTTCCCACGGGC
TTGGGGGAAAGATGGTGGGGACCAAGGCCTGGGTGTTCTCCTTCTGCTGGTCTGGAAGTCACA
TCTGTGTTGGGGAGACAGACGATGCTCACCAGTCAGTAAGAAGAGTCCAGCCTGGGAAGAA
GAACCCAGCATCTTTGCCAAGCCTGCCGACACCCTGGAGAGCCCTGGTGAAGTGGACAACAT
GGTTCAACATCGACTACCCAGGCGGGAAGGGCGACTATGAGCGGCTGGACGCCATTGCTTC
TACTATGGGGACCGTGTATGTGCCCGTCCCCTGCGGCTAGAGGCTCGGACCACCTGACTGGAC
ACCTGCGGGCAGCACTGGCCAGGTGGTCCATGGTAGTCCCCGTGAGGGTTTCTGGTGCCTCA
ACAGGGAGCAGCGGCCTGGCCAGAAGTCTCTAATTACACCGTACGCTTCTCTGCCACCA
GGATCCCTGCGCCGAGACACAGAGCGCATCTGGAGCCCATGGTCTCCCTGGAGCAAGTGCTC
AGCTGCCTGTGGTCAGACTGGGGTCCAGACTCGCACACGCATTGTCTGGCAGAGATGGTGT
CGCTGTGCAGTGAGGCCAGCGAAGAGGGTCAGCACTGCATGGGCCAGGACTGTACAGCCTGT
GACCTGACCTGCCCAATGGGCCAGGTGAATGCTGACTGTGATGCCTGCATGTGCCAGGACTT
CATGCTTCATGGGGCTGTCTCCCTTCCCGAGGTGCCCCAGCCTCAGGGGCTGCTATCTACC
TCCTGACCAAGACGCCGAAGCTGCTGACCCAGACAGACAGTGTGAGGAGATTCCGAATCCCT
GGCTTGTGCCCTGATGGCAAAAGCATCCTGAAGATCACAAGGTCAAGTTTGGCCCCATTGT
ACTCACAATGCCCAAGACTAGCCTGAAGGCAGCCACCATCAAGGCAGAGTTTGTGAGGGCAG
AGACTCCATACATGGTGATGAACCCTGAGACAAAAGCACGGAGAGCTGGGCAGAGCGTGTCT
CTGTGCTGTAAGGCCACAGGGAAGCCCAGGCCAGACAAGTATTTTGGTATCATAATGACAC
ATTGCTGGATCCTTCCCTCTACAAGCATGAGAGCAAGCTGGTGTGAGGAACTGCAGCAGC
ACCAGGCTGGGGAGTACTTTTGAAGGCCAGAGTGATGCTGGGGCTGTGAAGTCCAAGGTT
GCCCAGCTGATTGTACAGCATCTGATGAGACTCCTTGAACCCAGTTCTTGAGAGCTATCT
TATCCGGCTGCCCATGATTGCTTTGAGAATGCCAACCTCCTTCTACTATGACGTGGGAC
CCTGCCCTGTAAAGACTTGTGCAGGGCAGCAGGATAATGGGATCAGGTGCCGTGATGCTGTG
CAGAACTGCTGTGGCATCTCCAAGACAGAGGAAAGGGAGATCCAGTGCAGTGGCTACACGCT
ACCCACCAAGGTGGCCAAGGAGTGACGCTGCCAGCGGTGTACGGAACTCGGAGCATCGTGC
GGGGCCGTGTGAGTGTGCTGACAATGGGGAGCCCATGCGCTTTGGCCATGTGTACATGGGG
AACAGCCGTGTAAGCATGACTGGCTACAAGGGCACTTTACCCCTCCATGTCCCCCAGGACAC
TGAGAGGCTGGTGTGCTCACATTTGTGGACAGGCTGCAGAAGTTTGTCAACACCACCAAGTGC
TACCTTTCAACAAGAAGGGGAGTGCCGTGTTCCATGAAATCAAGATGCTTCGTGCGAAAGAG
CCCATCACTTTGGAAGCCATGGAGACCAACATCATCCCCCTGGGGGAAGTGGTTGGTGAAGA
CCCCATGGCTGAACTGGAGATTCCATCCAGGAGTTTCTACAGGCAGAATGGGGAGCCCTACA
TAGGAAAAGTGAAGGCCAGTGTGACCTTCTGGATCCCCGGAATATTTCCACAGCCACAGCT
GCCCAGACTGACCTGAACTTCATCAATGACGAAGGAGACACTTTCCCCCTTCGGACGTATGG
CATGTTCTCTGTGGAATTCAGAGATGAGGTCACTCAGAGCCACTTAATGCTGGCAAAGTGA
AGGTCACCTTGACTCGACCCAGGTCAAGATGCCAGACACATATCCACAGTGAACTCTGG
TCACTCAATCCAGACACAGGGCTGTGGGAGGGAAGGTGATTTCAAATTTGAAAATCAAAG
GAGGAACAAAAGAGAAGACAGAACCTTCTGGTGGGCAACCTGGAGATTCTGTGAGGAGGC
TCTTTAACCTGGATGTTCTGAAAGCAGGCGGTGCTTTGTTAAGGTGAGGGCTACCGGAGT
GAGAGGTTCTTGCTAGTGAGCAGATCCAGGGGTTGTGATCTCCGTGATTAACTGGAGCC
TAGAACTGGCTTCTGTCCAACCTAGGGCTGGGGCCGCTTTGACAGTGTATCACAGGCC
CCAACGGGGCCTGTGTGCCTGCCTTCTGTGATGACCAGTCCCCTGATGCCTACTCTGCCTAT
GTCTTGGCAAGCCTGGCTGGGGAGGAACTGCAAGCAGTGGAGTCTTCTCCTAAATTCAACCC
AAATGCAATTGGCGTCCCTCAGCCCTATCTCAACAAGCTCAACTACCGTCGGACGGACCATG
AGGATCCACGGGTTAAAAGACAGCTTTCCAGATTAGCATGGCCAAGCCAAGGCCCAACTCA
GCTGAGGAGAGCAATGGGCCCATCTATGCCTTTGAGAACCTCCGGGCATGTGAAGAGGCACC
ACCCAGTGACGCCACTTCCGGTCTACAGATTGAGGGGGATCGATATGACTACAACACAG
TCCCCTTCAACGAAGATGACCCTATGAGCTGGACTGAAGACTATCTGGCATGGTGGCCAAAG
CCGATGGAATTGAGGGCCTGCTATATCAAGGTGAAGATTGTGGGGCCACTGGAAGTGAATGT
GCGATCCCGCAACATGGGGGGCACTCATCGGCGGACAGTGGGGAAGCTGTATGGAATCCGAG
ATGTGAGGAGCACTCGGGACAGGACAGCCCAATGTCTCAGCTGCCTGTCTGGAGTTCAAG
TGCAGTGGGATGCTCTATGATCAGGACCGTGTGGACCGCACCCCTGGTGAAGGTCAATCCCCA
GGGCAGCTGCCGTGAGCCAGTGTGAACCCCATGCTGTCATGAGTACCTGGTCAACCACTGGC
CACTTGCACTCAACAACGACACCAGTGAGTACACCATGCTGGCACCCCTGGACCCCACTGGC
CACAACATATGGCATCTACACTGTCACTGACCAGGACCCCTCGCACGGGCCAAGGAGATCGCGCT
CGGCCGGTGCTTTGATGGCACATCCGATGGCTCCTCCAGAATCATGAAGAGCAATGTGGGAG
TAGCCCTCACCTTCAACTGTGTAGAGAGGCAAGTAGGCCGCCAGAGTGCCTTCCAGTACCTC
CAAAGCACCCAGCCAGTCCCCTGCTGCAGGCACTGTCCAAGGAAGAGTGCCTTCGAGGAG
GCAGCAGCGAGCGAGCAGGGGTGGCCAGCGCCAGGTTGGAGTGGTGGCCTCTCTGAGATTC

F4/310

FIGURE 71B

CTAGAGTTGCTCAACAGCCCCTGATCAACTAAAGTTTTGTGGTACTTCACCCCTCTTCTGCCCT
CATTTTCATGTGACAGCCATTGTGAGACTGATGCACAACTGTCACCTGGTTAATTTAAGCAC
TTCTGTTTTTCGTGAATTTGCTTGTTTGTTCCTTCATGCCTTTACTTACTTTGTCCCATGCTA
CTGATTGGCACGTGGCCCCCACAATGGCACAAATAAAGCCCCTTTGTGAAACTGTTCTTTAAA
TGAAACACAAGAAATTGGCCACTGGTAAACTCTGCAGCTTCAACTGTACTTCATTTAATGC
CATTAATGCAAATATACTTCCTCTTCTTTTTGCATGGTTTTGCCACCTCTGCAATAGTGAT
AATCTGATGCTGAAGATCAAATAACCAATATAAAGCATATTTCTTGGCCTTGCTCCACAGGA
CATAGGCAAGCCTTGATCATAGTTCATACATATAAATGGTGGTGAAATAAAGAAATAAAACA
CAATACTTTTACTTGAAATGTAAATACTTATTTATTTCTTGCTAAATTTGGAATTCTAGT
GCACATTCAAAGTTAAGCTATTAAATATAGGGTGATCATAGTTCCTCTACCAAGTCTGGAAA
GAACATCTCCTGGTATCCACAATTACACCAGGTTGCTAACTGTATTTGTACATTTCCCTTTG
CATTCGCTTTTGTCTTGCTAGAAACCCAGTGTAGCCAGGGCAGATGTCAATAAATGCATA
CTCTGTATTTTCGAAAAA

75/310

FIGURE 72

MVGTKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSIFAKPADTLESPGEWTTWFNI
DYPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQ
RPGQNCNSNYTVRFLCPPGSLRRDTERIWSPWSPWSKCSAACGQTGVQTRTRICLAEMVSLCS
EASEEGQHCMGQDCTACDLTCPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAYLLTK
TPKLLTQTDSDGRFRIPGLCPDGKSILKITKVKFAPIVLTMPKTSLKAATIKAEFVRAETPY
MVMNPETKARRAGQSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRLKQQHQAG
EYFCKAQSDAGAVKSKVAQLIVTASDETPCNPVPESYLIRLPHDCFQONATNSFYDVGRCPV
KTCAGQQDNGIRCRDAVQNC CGISKTEEREIQCSGYTLPTKVAKECSCQRCTETRSIVRGRV
SAADNGEPMRFGHVYMGNSRVSMGTGYKGTFTLHVPQDTERLVLT FVDRLQKFVN TTKVL PPN
KKGSAVFHEIKMLRRKEPITLEAMETNIIPLGEVVGEDPMAELEIPSR SFYRQNGEPYIGKV
KASVTFLDPRNISTATAAQTDLNFINDEGDTFPLRTYGMFSVDFRDEVTSEPLNAGKVKVHL
DSTQVKMPEHISTVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLFNL
DVPESRRRCFVKVRAYRSERFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGA
CVPAFCDDQSPDAY SAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLN YRRTDHEDPR
VKKTAFQISMAKPRPNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYNTVPFN
EDDPMSWTE DYLAWWPKPMEFRACYIKVKIVGPLEVNVR SRNMGGTHRRTVGKLYGIRDVRS
TRDRDQPNVSAACLEFKCSGMLYDQDRVDRTL VKVI PQGSCRRASVNPMLHEYLVNHLPLAV
NNDTSEYTMLAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALT
FNCVERQVGRQSAFQYLQSTPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFPRVA
QQPLIN

76/310

FIGURE 73

CTGCAAGTTGTTAACGCCTAACACACAAGTATGTTAGGCTTCCACCAAAGTCCTCAATATAC
CTGAATACGCACAATATCTTAACTCTTCATATTTGGTTTTGGGATCTGCTTTGAGGTCCCAT
CTTCATTTAAAAAAAATACAGAGACCTACCTACCCGTACGCATACATACATATGTGTATAT
ATATGTAACTAGACAAAGATCGCAGATCATAAAGCAAGCTCTGCTTTAGTTTCCAAGAAGA
TTACAAAGAATTTAGAGATGATTTTGTCAAGATCCCTGTGATTATGCCCTTTGGGTTACG
GTGTCCTCAGTGATGCAGCCCTACCCTTTGGTTTGGGGACATTATGATTTGTGTAAGACTCA
GATTTACACGGAAGAAGGAAAGTTTGGGATTACATGGCCTGCCAGCCGGAATCCACGGACA
TGACAAAATATCTGAAAGTGAAACTCGATCCTCCGGATATTACCTGTGGAGACCTCCTGAG
ACGTTCTGTGCAATGGGCAATCCCTACATGTGCAATAATGAGTGTGATGCGAGTACCCCTGA
GCTGGCACACCCCCCTGAGCTGATGTTTGAATTTTGAAGGAAGACATCCCTCCACATTTTGGC
AGTCTGCCACTTTGGAAGGAGTATCCCAAGCCTCTCCAGGTTAACATCACTCTGTCTTGAGC
AAAACCATTTAGCTAACAGACAACATAGTTATTACCTTTGAATCTGGGCGTCCAGACCAAAT
GATCCTGGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAG
ACTGCTTAGATGCTTTTACATGGATCCTAAATCCGTGAAGGATTTATCACAGCATACGGTC
TTAGAAATCATTTGCACAGAAGAGTACTCAACAGGGTATACAACAAATAGCAAATAATCCA
CTTTGAAATCAAAGACAGGTTGCGCTTTTGTCTGGACCTCGCCTACGCAATATGGCTTCCC
TCTACGGACAGCTGGATACAACCAAGAACTCAGAGATTTCTTTACAGTCACAGACCTGAGG
ATAAGGCTGTAAAGACCAGCCGTTGGGGAAATATTTGTAGATGAGCTACACTTGGCACGCTA
CTTTTACGCGATCTCAGACATAAAGGTGCGAGGAAGGTGCAAGTGTAATCTCCATGCCACTG
TATGTGTGTATGACAACAGCAAATTGACATGCGAATGTGAGCACAACTACAGGTCCAGAC
TGTGGGAAATGCAAGAAGAATTATCAGGGCCGACCTTGGAGTCCAGGCTCCTATCTCCCCAT
CCCCAAAGGCACTGCAAATACCTGTATCCCCAGTATTTCCAGTATTGGTACGAATGTCTGCG
ACAACGAGCTCCTGCACTGCCAGAACGGAGGGACGTGCCACAACAACGTGCGCTGCCTGTGC
CCGGCCGCATACACGGGCATCCTCTGCGAGAAGCTGCGGTGCGAGGAGGCTGGCAGCTGCGG
CTCCGACTCTGGCCAGGGCGCGCCCCCGCACGGCACCCAGCGCTGCTGCTGCTGACCACGC
TGCTGGGAACCGCCAGCCCCCTGGTGTCTAGGTGTACCTCCAGCCACACCGGACGGGCCT
CTGCCGTGGGGAAGCAGACACAACCCAAACATTTGCTACTAACATAGGAAACACACATAC
AGACACCCCCACTCAGACAGTGTACAACTAAGAAGGCCTAACTGAACCTAAGCCATATTTAT
CACCCGTGGACAGCACATCCGAGTCAAGACTGTTAATTTCTGACTCCAGAGGAGTTGGCAGC
TGTTGATATTATCACTGCAAATCACATTGCCAGCTGCAGAGCATATTGTGGATTGGAAAGGC
TGCGACAGCCCCCAAACAGGAAAGACAAAAAACAACAAATCAACCGACCTAAAAACATTG
GCTACTCTAGCGTGGTGCGCCCTAGTACGACTCCGCCCAGTGTGTGGACCAACCAAATAGCA
TTCTTTGCTGTGAGGTGCATTGTGGGCATAAGGAAATCTGTTACAAGCTGCCATATTGGCCT
GCTTCCGTCCCTGAATCCCTTCCAACCTGTGCTTTAGTGAACGTTGCTCTGTAACCTCGTT
GGTTGAAAGATTTCTTTGTCTGATGTTAGTGATGCACATGTGTAACAGCCCCCTCTAAAAGC
GCAAGCCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCGAGCACACCCAC
TATACAAGAGTGGCTATAGGAAAAAAGAAAGTGTATCTATCCTTTTGTATTCAAATGAAGTT
ATTTTTCTTGAACCTACTGTAATATGTAGATTTTTTGTATTATTGCCAATTTGTGTTACCAGA
CAATCTGTTAATGTATCTAATTCGAATCAGCAAAGACTGACATTTTATTTTGTCTCTTTG
TTCTGTTTTGTTTTCACTGTGCAGAGATTTCTCTGTAAGGGCAACGAACGTGCTGGCATCAA
GAATATCAGTTTACATATATAACAAGTGTAATAAGATTCCACCAAAGGACATTCTAAATGTT
TTCTTGTTGCTTTAACTGGAAGATTTAAGAATAAAAACTCCTGCATAAACGATTTTCAAG
AATTTGTATTGCAATTTCTTAAGATGAAAGGAACAGCCACCAAGCAGTTTCACTCACTTT
ACTGATTTCTGTGTGGACTGAGTACATTGAGCTGACGAATTTAGTTCCAGGAAGATGGATT
GATGTTCACTAGCTTGGACAACTTCTGCAAAATATGAGACTATTTCCACTTGGGAAAAATTA
CAACAGCAAAAAAAAAAAAAAAAAAAAAA

77/310

FIGURE 74

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLK
VKLDPPDITCGDPPETFCAMGNPYMCNNECDASTPELAHPPELMFDFEGRHPSTFWQSATWK
EYPKPLQVNITLSWSKTIELTDNIVITFESGRPDQMILEKSLDYGRTWQPYQYYATDCLDAF
HMDPKSVKDLSQHTVLEI ICTEEYSTGYTTNSKI IHFEIKDRFALFAGPRLRNMASLYGQLD
TTKKLRDFFFTVTDLRIRLLRPAVGEIFVDELHLARYFYAISDIKVRGRCKCNLHATVCVYDN
SKLTCECEHNTTGPDCGKCKKNYQGRPWSPGSYLPI PKGTANTCIPSISSIGTNVCDNELLH
CQNGGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGQGAPPHGTPALLLLTTLLGTAS
PLVF

78/310

FIGURE 75

CCCACGCGTCCGGGTGACCTGGGCCGAGCCCTCCCGGTGGCTAAGATTGCTGAGGAGGCGG
CGGGTAGCTGGCAGGCGCCGACTTCCGAAGGCCGCGGTCCGGGCGAGGTGTCTCATGACTT
CTCTTGTTGGACCATGTCCTGTATCTTTTTTGCCTGCGTGGTACGGGTAAGGGATGGACTGCC
CCTCTCAGCCTCTACTGATTTTTTACCACACCCAAGATTTTTTGGAAATGGAGGAGACGGCTCA
AGAGTTTAGCCTTGCGACTGGCCCAGTATCCAGGTCGAGGTTCTGCAGAAGGTTGTGACTTT
AGTATACATTTTTCTTCTTTTGGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCC
AGCAGCCATGGCCTTCTGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCCTATGACA
CTACCTGCATTGGCCTAGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTGAG
AAAGTGAAGTGGCATTTTAACTATGTAAGTTCCTCTCAGATGGAGTGCAGCTTGGAATAAT
TCAAGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCTGGAGGACACAGATGTGGCAA
ATGGGGTGATGAATGGTCACACACCGATGCACTTGGAGCCTGCTCCTAATTTCCGAATGGAA
CCAGTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCT
CATTCGAGGAGTTACCTTGCGAAGCATTCTTTACAGGATCCAAGGAGCTGGTTCTGCTGGT
TGGACCAAACCTCGTGAGGCCAGCCACCCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCAT
CCGGGAGCAGTGATGTCAAACCTTCTGCTGCTGGGGAAATCTCATCAGCAGGGAGCCTGTGGA
AAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTCCGGAACATCTGCCCATGTGTATTG
ATGGCAGAGCTGTTGCCCCACAAGCGCCTTTTATTTAGGGTAAAATTAACAAATCCATTCTAT
TCCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTT
GCTTCAGAAGTGTTATTTTCATGAATCATTCATATGATTTGATCCCCCAGGATTCTATTTTGT
TTAATGGGCTTTTCTACTAAAAGCATAAAATACTGAGGCTGATTTAGTCAGGGCAAACCAT
TTACTTTACATATTCGTTTTCAATACTTGCTGTTTCATGTTACACAAGCTTCTTACGGTTTTTC
TTGTAACAATAAATATTTTGAGTAAATAATGGGTACATTTTAACAACTCAGTAGTACAACC
TAACTTGTATAAAAGTGTGTAAAAATGTATAGCCATTTATATCCTATGTATAAATTAATG
AGGTGGCTTCAGAAATGGCAGAATAAATCTAAAGTGTATTATTAATAAAAAAAAAAAAAAAAAA
AAAAG

79/310

FIGURE 76

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFSIHF
SSFGDVACMAICSCQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQKVW
HFNYVSSSQMECSLEKIQEELKLQPPAVLTLEDTDVANGVMNGHTPMHLEPAPNFRMEPVTA
LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

FIGURE 77

TGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCNTATGACACTACCTGCATTGGCNT
AGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTCAGAAAGTGAAGTGGCATT
TTAACTATGTAAGTTCCTNTCAGATGGAGTGCAGCTTGGA AAAAATTCAGGAGGAGCTCAAG
TTGCAGCCTCCAGCGGTTCTCANTATGGAGGACACAGATGTGGCAAATGGGGT

81/310
FIGURE 78

CTCAGCGGCGCTTCCTCGTAGCGAGCCTAGTGCGGGGTGTTTGCATTGAAACGTGAGCGCGA
CCCCACCTTAAAGAGTGGGGAGCAAAGGGAGGACAGAGCCCTTTAAAACGAGGCGGGTGGTG
CCTGCCCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCGAGTT
TCTGTGCGCAGGCTGCGAGGAAAGGCCCTAGGCTGGGTCTGGGTGCTTGGCGGCGGCGGCTT
CCTCCCCGCTCGTCCTCCCCGGGCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTA
TGGAAGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCACGAGAGGATCCGC
GAGTGTATTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGAC
CCGCTTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGA
TTGCGCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCTCTGCTCCTGCCC
TTCTCCATCATCAGCAATGAGGTGCTGCTCTCCCTGCCTCGGAATACTACATCCAGTGGCT
CAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTTTTCTCTTCCCCAACCTGTCCCTCA
TCTTCCTCATGCCCTTTGCATATTTCTTCACTGAGTCTGAGGGCTTTGCTGGCTCCAGAAAG
GGTGTCTGGGCCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCT
AGGTATGGTGTGGGTGGCATCAGCCATTGTGGACAAGAACAAGGCCAACAGAGAGTCACTCT
ATGACTTTTGGGAGTACTATCTCCCTACCTCTACTCATGCATCTCCTTCCTTGGGGTTCTG
CTGCTCCTGGTGTGTACTCCACTGGGTCTCGCCCGCATGTTCTCCGTCACTGGGAAGCTGCT
AGTCAAGCCCCGGCTGCTGGAAGACCTGGAGGAGCAGCTGTACTGCTCAGCCTTTGAGGAGG
CAGCCCTGACCCGAGGATCTGTAATCCTACTTCCTGCTGGCTGCCTTTAGACATGGAGCTG
CTACACAGACAGGTCCTGGCTCTGCAGACACAGAGGGTCTGCTGGAGAAGAGGCGGAAGGC
TTCAGCCTGGCAACGGAACCTGGGCTACCCCTGGCTATGCTGTGCTTGCTGGTGCTGACGG
GCCTGTCTGTGCTCATTGTGGCCATCCACATCCTGGAGCTGCTCATCGATGAGGCTGCCATG
CCCCGAGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCCTTCTCCAAGCTGGGCTCCTTTGG
TGCCGTCATTAGGTTGTACTCATCTTTTACCTAATGGTGTCTCCTCAGTTGTGGGCTTCTATA
GCTCTCCACTCTTCCGGAGCCTGCGGCCAGATGGCACGACACTGCCATGACGCAGATAATT
GGGAACTGTGTCTGTCTCCTGGTCCTAAGCTCAGCACTTCCTGTCTTCTCTCGAACCCTGGG
GCTCACTCGCTTTGACCTGCTGGGTGACTTTGGACGCTTCAACTGGCTGGGCAATTTCTACA
TTGTGTTCTCTACAACGCAGCCTTTGCAGGCCTCACCACACTCTGTCTGGTGAAGACCTTC
ACTGCAGCTGTGCGGGCAGAGCTGATCCGGGCCTTTGGGCTGGACAGACTGCCGCTGCCCGT
CTCCGGTTTCCCCCAGGCATCTAGGAAGACCCAGCACCACTGAGCCTCCAGCTGGGGGTGGGA
AGGAAAAAACTGGACACTGCCATCTGCTGCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGG
ACCTCAGGACCTGGAATCTGAGAGGGTGGGTGGCAGAGGGGAGCAGAGCCATCTGCACTATT
GCATAATCTGAGCCAGAGTTTGGGACCAGGACCTCCTGCTTTTCCATACTTAAGTGTGGCCT
CAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCCTGATCCCAAATCTGTTTACACATCA
ATCTGCCTCACTGCTGTTCTGGGCCATCCCCATAGCCATGTTTACATGATTTGATGTGCAAT
AGGGTGGGGTAGGGGCAGGGAAAGGACTGGGCCAGGGCAGGCTCGGGAGATAGATTGTCTCC
CTTGCTCTGGCCCAGCAGAGCCTAAGCACTGTGCTATCCTGGAGGGGCTTTGGACCACCTG
AAAGACCAAGGGGATAGGGAGGAGGAGGCTTCAGCCATCAGCAATAAAGTTGATCCCAGGGA
AAAAAA

82/310

FIGURE 79

MEAPDYEVLSVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDDDEDATVNK
IALELCTFTLAIALGAVLLLPFSIISNEVLLSLPRNYYIQWLNGSLIHGLWNLVFLFPNLSL
IFLMPFAYFFTESEGFAGSRKGV LGRVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL
YDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTGKLLVKPRLLEDLEEQLYCSAFEE
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRV LLEKRRKASAWQRNLGYPLAMLCLLVLT
GLSVLIVAIHILELLIDEAAMP RGMQGTSLGQVSFSKLGSFGAVIQVVLIFYLMVSSVVG FY
SSPLFRSLRPRWHD TMTQIIGNCVCLLV LSSALPVFSRTLGLTRFDLLGDFGRFNWLG NFY
IVFLYNAAFAGLTTLC LVKTFTA AVRAELIRAFGLDRLPLPVSGFPQASRKTQH Q

83/310

FIGURE 80

GGCTGCCGAGGGAAGGCCCTTGGGTTGGTCTTGGTTGCTTGGCGGCGGCGGNTTCNTCCCC
GCTCGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGAAGC
ACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGTGTA
TTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGACCCGCTTC
AAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCG

FIGURE 81

GACCGACCTTAAAGAGTGGGAGCAAAGGGAGGACAGAGCCTTTTAAAACGAGGCGGTGGTG
CTGCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTTTC
TGTCGCAGGCTGCGAGGAAAGGCCCTAGGCTGGGTCTGGTGCTTGGCGGCGGCGGCTTCCT
CCCCGTTGTCNTCCCCGGGCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGA
AGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGT
GTATTATATCAACACTTCTGTTTGCAACACTGTACATCNTCTGCCACATCTTCCTGACCCGC
TTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGATTGC
GCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCCTGCTCCTGCCCTTCT
CCATCATCAGCAATGAGGTGCTGCACTCCC

FIGURE 82

GATGTGCTCCTTGGAGCTGGTGTGCAGTGTCTGACTGTAAGATCAAGTCCAAACCTGTTTT
GGAATTGAGGAACTTCTCTTTTGATCTCAGCCCTTGGTGGTCCAGGTCTTCATGCTGCTGT
GGGTGATATTACTGGTCCTGGCTCCTGTCAGTGGACAGTTTGCAAGGACACCCAGGCCCAT
ATTTTCCTCCAGCCTCCATGGACCACAGTCTTCCAAGGAGAGAGAGTGACCCTCACTTGCAA
GGGATTTGCTTCTACTCACCACAGAAAACAAATGGTACCATCGGTACCTTGGGAAAGAAA
TACTAAGAGAAACCCAGACAATATCCTTGAGGTTTCAAGGAATCTGGAGAGTACAGATGCCAG
GCCCAGGGCTCCCCTCTCAGTAGCCCTGTGCACTTGGATTTTTCTTCAGAGATGGGATTTCC
TCATGCTGCCCAGGCTAATGTTGAACTCCTGGGCTCAAGTGATCTGCTCACCTTAGGCCTCTC
AAAGCGCTGGGATTACAGCTTCGCTGATCCTGCAAGCTCCACTTTCTGTGTTTGAAGGAGAC
TCTGTGGTTCTGAGGTGCCGGGCAAAGGCGGAAGTAACACTGAATAATACTATTTACAAGAA
TGATAATGTCCTGGCATTCTTAATAAAAGAACTGACTTCCAAAAAAAAAAAAAAAAAAAAA

FIGURE 83

MLLWVILLVLAPVSGQFARTPRPIIFLQPPWTTVFQGERVTLTCKGFRFYSPQKTKWYHRYL
GKEILRETPDNILEVQESGEYRCQAQGSPLSSPVHLD FSSEMGFPHAAQANVELLGSSDLLT

FIGURE 84

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCCCGGTGT
GAGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTGTGAG
GAGTGTGTGGAACAGGACCCGGGACAGAGGAACCATGGCTCCGCAGAACCTGAGCACCTTTT
GCCTGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGACGAGATTTCTATAAGATCTTG
GGGGTGCCTCGAAGTGCCTCTATAAAGGATATTA AAAAGGCCCTATAGGAACTAGCCCTGCA
GCTTCATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCAGGATCTGGGTG
CTGCTTATGAGGTTCTGTCAGATAGTGAGAAACGGAAACAGTACGATACTTATGGTGAAGAA
GGATTAAAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTTCACACTTCTTTGGGGATTT
TGGTTTCATGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCCAAGAGGAAGTGATA
TTATGTAGATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTT
AGAAACAAACCTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTCGGCAAGAGAT
GCGGACCACCCAGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAAT
GCCCTAATGTCAAACCTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTG
AGAGACGGCATGGAGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGG
AGATTTACGGTTCCGAATCAAAGTTGTCAAGCACCCAATATTTGAAAGGAGAGGAGATGATT
TGTACACAAATGTGACAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACT
CACTTGGATGGTCACAAGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCT
ATGGAAGAAAGGGGAAGGGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAA
TCACTTTTGATGTGGATTTTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAA
CAGCTACTGAAACAAGGGTCAGTGCAGAAGGTATACAATGGACTGCAAGGATATTGAGAGTG
AATAAAATTGGACTTTGTTTAAAATAAGTGAATAAGCGATATTTATTATCTGCAAGGTTTTT
TTGTGTGTGTTTTTGTTTTTATTTTCAATATGCAAGTTAGGCTTAATTTTTTTATCTAATGA
TCATCATGAAATGAATAAGAGGGCTTAAGAATTTGTCCATTTGCATTTCGGAAAAGAATGACC
AGCAAAAGGTTTACTAATACCTCTCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGT
TTCAAGAATTAAAGCTGCAAGAGGACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGA
GTTGTTAGCAATTTTATTCAAATGCCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTG
TTATTTTTTA

FIGURE 86

TGGGACCAGGGAACCCCGGGCCCCCGGTGGAGNGCCTAACAGGCCGGTGGNTGCGACCGAA
GCGGCGGGCGGAGGAGGTTTTGAGGATTTTTGGAACAGGACCCGGACAGAGGAACCATGGTT
CCGCAGAACNTGAGCACNTTTTGCCTGTTGNTGNTATACTTCATCGGGGCGGTGATTGCCGG
ACGAGATTTNTATAAGATTTTGGGGTGCCTNGAAGTGCCTTNTATAAAGGATATTAAAAAGG
CCTATAGGAAACTAGCCCTGCAGNTTTATCCCGACCGGAACCCCTGATGATCCACAAGCCCAG
GAGAAATTCCAGGATTTGGGTGCTGCTTATGAGGTTNTGTCAGATAGTGAGAAACGGAAACA
GTACGATAATTATGGTGAAGAAGGATTAAAAGATGGTNATCAGAGCTCCCATGGAGACATTT
TTTCACACTTNTTTGGGGATTTTGGTTTCATGTTTGGAGGAACCCCTNGTCAGCAAGACAGA
AATATTCCAAGAG

20/310

FIGURE 87

GGCACGAGGCGGCGGGGCAGTCGCGGGATGCGCCCGGGAGCCACAGCCTGAGGCCCTCAGGT
CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA
GCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAAGCATGG
AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGTTTGCAGCC
TTGGTGCTGGTTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGCGCTATGATTC
TAAGCCCATTGTGGACCTCATTTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC
TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGAATGAAGACTGGATC
GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC
AGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA
GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG
TACCCTCCGTTGGACCCCAAACCTCCTGGACGCACGGACGACTGCCCTGCTCCTGTCTGTCAG
TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC
AGTCTCTGTGCGCTGCTGAGGAGCATTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG
CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATTTTAGTGCCT
ACAGGCCAGCAGCTAGCCATGAAGGCCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT
CTACTTTTTCTATAGAGTTAGTTGTTCTCCACGGCTGGAGAGTTCAGCTGTGTGTGCATAG
TAAAGCAGGAGATCCCCGTCAGTTTATGCCTCTTTTGCAGTTGCAAACCTGTGGCTGGTGAGT
GGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCTGCAAGAGGAGTATTGAAAA
CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCT
AAGAAATCAAGAGGTTTACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTCAGAATG
TGTGGCAATTCTGATCTGCATTTTCAGAAGAGGACAATCAATTGAACTAAGTAGGGGTTTC
TTCTTTTGGCAAGACTTGTA CTCTCACCTGGCCTGTTTCATTTATTTGTATTATCTGCCT
GGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGCAAGTTTGGGTTTGAAGCTGAGGAACT
ACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTG
GATAGTAAATTTATACTTATGTTTCCCTCAAAAAAAAAAAAAA

FIGURE 88

METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSEL
ELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSAS
VSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDARTTALLSVSHLVLVTRNACHLTGGLDWI
DQSLSAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

28/310

FIGURE 89

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTCCGAGGTGCTTTCGCCGCTGTCC
CCACCACTGCAGCCATGATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA
TTTGGAGTGTTTTCTCTGTTCTTTGGAATGATTCTCTTTTTTGACAAAGCACTACTGGCTAT
TGGAAATGTTTTATTTGTAGCCGGCTTGGCTTTTGTAAATTGGTTTAGAAAGAACATTCAGAT
TCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTTCTGGGTGGTGTATTTGTAGTC
CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTCTCTTGTTTCAG
GGGCTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGGATCCCTCCTAAAT
TTACCTGGAATTAGATCATTTGTAGATAAAGTTGGAGAAAGCAACAATATGGTATAACAACA
AGTGAATTTGAAGACTCATTTAAAATATTGTGTTATTTATAAAGTCATTTGAAGAATATTCA
GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAAACGTATAG
CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA
ACTAAGAAGAAGTCAGCAAGCAAACCTGAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAA
ACTCTTGAAGGCTATTTGTGTTGTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA
CTGTGGTGCCTGTTTCTTTCTTTTATTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT
TTTTAGAAGTGTCCTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA
TGCATGAATTCGATTGGATTGTGTCATTTTAAAGTATTAAAACCAAGGAAACCCCAATTTTG
ATGTATGGATTACTTTTTTTTGNGCNCAGGGCC

FIGURE 90

MISLTDQTQKIGMGLTGFGVFFLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQK
HKMKATGFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFI RRVPVLGSLLNLPGI
RSFVDKVGESNNMV

Important features:**Transmembrane domains:**

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

N-myristoylation sites.

amino acids 11-16, 51-56 and 116-121

Aminoacyl-transfer RNA synthetases class-II protein.

amino acids 49-59

94/310

FIGURE 91

GAAGACGTGGCGGCTCTCGCCTGGGCTGTTTCCCGGCTTCATTTCTCCCGACTCAGCTTCCC
ACCNTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCCCCACCACTGCAGCCATGATCTCCTTAA
CGGACACGCAGAAAATTGGAATGGGATTAACCGGATTTGGAGTGTTTTCTGTTCCTTGA
ATGATTCTCTTTTTTGACAAAGCACTACTGGCTATTGGAAATGTTTTATTTGTAGCCGGCTT
GGCTTTTGTAATTGGTTTAGAAAGAACATTCAGATTCTTCTTCCAAAAACATAAAATGAAAG
CTACAGGTTTTTTTTCTGGGTGGTGTATTTGTAGTCCTTATTGGTTGGCCTTTGATAGGCATG
ATCTTCGAAATTTATGGATTTTTTCTCTTGTTT

25/310

FIGURE 92

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAA
GGCTGCCAGGAAGGAGACGCCTTCCTGAGTCTTGATCTTTCTTCCTTCTGGAAATCTTTGA
CTGTGGGTAGTTATTTATTTCTGAATAAGAGCGTCCACGCATCATGGACCTCGCGGGACTGC
TGAAGTCTCAGTTCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATC
ATCAACACCATTAGCTCTTCACTCTCCTCCTCTGGCCCATTAACAAGCAGCTCTTCCGGAA
GATCAACTGCAGACTGTCCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGT
CGGGCACGGAATGCACCATCTTCACGGACCCGCGCGCCTACCTCAAGTATGGGAAGGAAAAT
GCCATCGTGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGCTGGAGCCTGTCCGA
ACGCTTTGGGCTGTTAGGGGGCTCCAAGGTCTGGCCAAGAAAGAGCTGGCCTATGTCCCAA
TTATCGGCTGGATGTGGTACTTCACCGAGATGGTCTTCTGTTCGCGCAAGTGGGAGCAGGAT
CGCAAGACGGTTGCCACCAGTTTGACGACCTCCGGGACTACCCCGAGAAGTATTTTTTCTCT
GATTCACTGTGAGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCC
GGGCAAGGGGCTGCCTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATC
ACCGTGAGGAGCTTGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTTAGAAA
TAATGAAAATCCAACACTGCTGGGAGTCTAAACGGAAAGAAATACCATGCAGATTTGTATG
TTAGGAGGATCCCACTGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCAC
AAGCTCTACCAGGAGAAGGATGCCTTTTCAAGGAGGAGTACTACAGGACGGGCACCTTCCCAGA
GACGCCCATGGTGCCCCCCCCGGCGGCCCTGGACCCTCGTGAACTGGCTGTTTTGGGCCTCGC
TGGTGCTCTACCCTTTCTTCCAGTTCCTGGTCAGCATGATCAGGAGCGGGTCTTCCCTGACG
CTGGCCAGCTTCATCCTCGTCTTCTTTGTGGCCTCCGTGGGAGTTCGATGGATGATTGGTGT
GACGGAAATTGACAAGGGCTCTGCCTACGGCAACTCTGACAGCAAGCAGAACTGAATGACT
GACTCAGGGAGGTGTCACCATCCGAAGGGAACCTTGGGGAACCTGGTGGCCTCTGCATATCCT
CCTTAGTGGGACACGGTGACAAAGGCTGGGTGAGCCCCTGCTGGGCACGGCGGAAGTCACGA
CCTCTCCAGCCAGGGAGTCTGGTCTCAAGGCCGGATGGGGAGGAAGATGTTTTGTAATCTTT
TTTTCCCATGTGCTTTAGTGGGCTTTGGTTTTCTTTTGTGCGAGTGTGTGTGAGAATGGC
TGTGTGGTGAGTGTGAACTTTGTTCTGTGATCATAGAAAGGGTATTTTAGGCTGCAGGGGAG
GGCAGGGCTGGGGACCGAAGGGGACAAGTTCCCCTTTCATCCTTTGGTGCTGAGTTTTCTGT
AACCCTTGTTGCCAGAGATAAAGTGAAAAGTGCTTTAGGTGAGATGACTAAATTATGCCTC
CAAGAAAAAAAATTAAAGTGCTTTTCTGGGTCAAAAAAAAAA

FIGURE 93

MDLAGLLKSQFLCHLVFCYVFIA SGLIINTIQLFTLLLWPINKQLFRKINCRLSYCISSQLV
MLEWWSGTECTIFTDPRAYLKYGKENAIVVLN HKFEIDFLCGWSLSERFGLLGGSKVLAKK
ELAYVPIIGWMWYFTEMVFC SRKWEQDRKTVATSLQH LRDYPEKYFFLIHCEGTRFTEKKHE
ISMQVARAKGLPRLKHLLPRTKGFAITVRSLRNVVSAVDCTLNFRNNENPTLLGV LNKKK
YHADLYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFPETPMVPPRRPWTLVN
WLFWASLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDS
KQKLND

FIGURE 94

CTGAGGCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCG
GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA
GTAAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA
TACAATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAG
GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT
TGGTACAAATTCCGTCGTCATTGAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAA
AACTTGCAGGAGCATTTTTCAAACCAAGACCTTGTTTTCTGCTATTAACACCAAGTATAA
TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTT
TTTCACAGGGTACCTTTAGTGGTTGCCAATCTGGGCATGTCTGAACAACTGGGTATATAAAC
TGTATCAGGTTCTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAAT
TTTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA
CAAGAGGAATTAAGAGTATATGCAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT
AGTAAAGGATGTAAACAGATTAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGG
CAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTA
CGGACCTTTTTTCCAAATTCTGAATTTCTTCATTGATGTTATGTCTTTAAAAAATAGACA
TGTTTTCTAAAAGTAGCTGTAACACCAACCCTCGATGTAGTAGACAATCTGACCTTAA
TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCAT
AAAGCCTTAGACTTAGATGACAGATGGCAATTCAGAGATCTCGGTTGTTAGATACACAAGA
CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCC
CAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACA
TTTTGATCCTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAACAT
TTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTAC
CTGTTTGCAGTAATACACAGATAACTCTTAGTGCAATTTACTTCACAAAGTACTTTTTCAAAC
ATCAGATGCTTTTATTTCCAAACCTTTTTTTTACCTTTCACTAAGTTGTTGAGGGGAAGGCT
TACACAGACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA
TCCCAGCACTTAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCC
TGGGCAACGTATTGAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTAT
TTTCAAATATGGAAAGAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAG
TGATACTTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCA
ATAAATTTGCAAAACATCATCTAAAATTTAAAAAATAAAAAAAAAAAAAAAAAAAAAA

FIGURE 95

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGGEVKGAKNSITDSQMDDVEVVYTIDIQ
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH
FSNQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC
MSTGFSRAVQTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN
RLKREIEKRRGAQIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS
CNYNHHLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDDRWFKRSRLDQTQDKRSKA
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

FIGURE 96

GGCACAGCCGCGCGGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCA
GCGCAGGGCAGCCCAAGCAGCGCGCAGCGAACGCCCGCCGCCACACCCTCTGCGGTCC
CCGCGGCGCCTGCCACCCTTCCCTCCTTCCCCGCGTCCCCGCTCGCCGGCCAGTCAGCTTG
CCGGGTTGCTGCCCCGCGAAACCCCGAGGTCACCAGCCCGCGCCTCTGCTTCCCTGGGCCG
CGCGCCGCTCCACGCCCTCCTTCTCCCCTGGCCCGGCGCCTGGCACCCGGGGACCGTTGCCT
GACGCGAGGCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCTGCTCGCCTCTTCCACCA
ACTCCAACCTCCTTCTCCCTCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCC
GCTGCCGTAGCGCCGCTTCCCGTCCGGTCCCAAAGGTGGGAACGCGTCCGCCCGGCCCCGCA
CCATGGCACGGTTTCGGCTTGCCCGCGCTTCTCTGCACCCTGGCAGTGCTCAGCGCCGCGCTG
CTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGACGTCTTTACGTGTCCAAAGG
CTTCAACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTGAAGATCTGTCCCC
AGGGTTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGAT
TTCAAAGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTACGTTACAA
GAAGTTTGATGAATTCTTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTGAATGATATGT
TTGTGAAGACATATGGCCATTTATACATGCAAAATTCTGAGCTATTTAAAGATCTCTTCGTA
GAGTTGAAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTG
GGCTCGCCTCCTGGAGCGGATGTTCCGCTGGTGAACCTCCAGTACCACTTTACAGATGAGT
ATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCTCGCAA
TTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTGCTCAAGGCTTAGCGGT
TGCGGGAGATGTCGTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGTAACCATGCCC
TGTTGAAGATGATCTACTGCTCCCACTGCCGGGGTCTCGTGACTGTGAAGCCATGTTACAAC
TACTGCTCAAACATCATGAGAGGCTGTTTGGCCAACCAAGGGGATCTCGATTTTGAATGGAA
CAATTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGAGGGTCCCTTCAACATTGAAT
CGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGATAATAGT
GTTCAAGTGTCTCAGAAGGTTTTCCAGGGATGTGGACCCCCCAAGCCCTCCAGCTGGACG
AATTTCTCGTTCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATCACCCGAGG
AACGCCCAACCACAGCAGCTGGCACTAGTTTGGACCGACTGGTTACTGATGTCAAGGAGAAA
CTGAAACAGGCCAAGAAATTCTGGTCCCTCCCTCCGAGCAACGTTTGCAACGATGAGAGGAT
GGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTACCTGT
TTGCAGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCAGGTTGACACC
AGCAAACCAGACATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGAT
GAAGAATGCATACAATGGGAACGACGTGGACTTCTTTGATATCAGTGATGAAAGTAGTGGAG
AAGGAAGTGGAAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAGTTTGACTACAATGCCACT
GACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGTCTCGGGCACA
GGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCCTGGTTATGCAGAGAGAGTGAGAGATAAT
TCTCAAACCTCTGAGAAAAAGTGTTTCATCAAAAAGTTAAAGGCACCAGTTATCACTTTTCTA
CCATCCTAGTGACTTTTGCTTTTTTAAATGAATGGACAACAATGTACAGTTTTTACTATGTGGC
CACTGGTTTAAAGAAGTGCTGACTTTGTTTTCTCATTGAGTTTGGGAGGAAAAGGGACTGTG
CATTGAGTTGGTTCCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGTGATGTTTGTATGTTTTTCTC
ATTTGCTTTGTGGGTTTTTTTTTCCAACGTGATCTCGCCTGTTTCTTACAAGCAAACCAG
GGTCCCTTCTTGGCACGTAACATGTACGTATTTCTGAAATATTAAATAGCTGTACAGAAGCA
GGTTTTATTTATCATGTTATCTTATTAAAGAAAAAGCCCCAAAAGC

FIGURE 97

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRRLYVSKGFNKNDAPLHEINGDHLKICPQ
GSTCCSQEMEKEYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMF
VKTYGHLVMQNSELFKDLFVELKRYVVGNNLEEMLNDFWARLLERMFLVNSQYHFTDEY
LECVSKYTEQLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHAL
LKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIES
VMDPIDVKISDAIMNMQDNSVQVSQKVFQGC GPPKPLPAGRISRSISESAFSARFRPHHPEE
RPTTAAGTSLDRLVTDVKEKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLF
AVTGNGLANQGNNPEVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGE
GSGSGCEYQQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

FIGURE 98

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCT
GACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAA
GCAACTTACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGC
TGCCACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCAC
AGGGACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAA
AGATTGGTTCCTGAGAGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGC
AGTGCCCCCTGTGATCATTTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCCACAGA
AAGCCAAACAAGCATTCCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTT
TGCTCTGCCTTTGTAGGAGCTCTGAGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAA
GACAGTGAGCACACCTACCAGACACTCTTCTTCTCCACCTCACTCTCCCACTGTACCCACC
CCTAAATCATTCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTTTGTTGCTCTC
TCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTT
AATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCTAGCTAGTGTCAATTTAACCTTAAATGC
AATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTTAAATGTCAAAAAAAAAAAAAAAAAA

-102/310

FIGURE 99

MKVLISSLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRR
KFMTVSGLPKKQCPDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

103/310

FIGURE 100

AATGGCTGTCTTAGTACTTCGCCTGACAGTTGTCCTGGGACTGCTTGTCTTATTCCTGACCT
GCTATGCAGACGACAAACCAGACAAGCCAGACGACAAGCCAGACGACTCGGGCAAAGACCCA
AAGCCAGACTTCCCCAAATTCCTAAGCCTCCTGGGCACAGAGATCATTGAGAATGCAGTCGA
GTTTCATCCTCCGCTCCATGTCCAGGAGCACAGGATTTATGGAATTTGATGATAATGAAGGAA
AACATTTCATCAAAGTGACATCCTCAGGACACACCCATGTGGCTCCTGGACAATCCAAGAGCA
GCCAAATCCTGCTTTTCCAGTTTGGCTCCACAAGTCCTCCAGGACAGAGCCCTCAAAGCAAC
TCCCAACGAGTTCTCAGGATTCAGGCTCTGGCTTCAACCAAACAGAACTCATTTTGAACACC
CTGACTGCATTTTTGCTTTTAGAAAGTTAGAATAAATATGGCGCTTTGGGATCACATAGTTG
ATGGAGAGGAAA

104/310

FIGURE 101

MAVLVLRRLTVVLGLLVFLTCYADDKPKPDDKPDDSGKDPKPDFPKFLSLLGTEIIENAVE
FILRSMRSTGFMEFDDNEGKHSSK

105/310

FIGURE 102

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT
CAGAGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTAC
CCTGCCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAGCTACTTCC
CCTACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAG
CTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG
CTGCGGAACCGGAGCCAACTTTTCACTTCTACCCACCGGGCTGCAGGGTCACTGCCTAGACC
CAAATCCCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT
GAGCGGTTTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGT
GGTGGTCTGCACTCTGGTGCTGTGCTCTGTGCAGAGCCCAAGGAAGGTCCTGCAGGAGGTCC
GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA
AGCTGGGCCTTCATGTGGCAGCAAGTTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTG
CTGCCTCACCAGAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGG
AACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC
AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA
AGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTAGCAGAATGAGAGAAGACATT
CATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC
CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC
TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC
CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT
CTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCCTGACCCTCT
CTCCCCACTACCACCTTCTTCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT
GCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGA
AACCACG

105/310

FIGURE 103

MDILVPLLQLLVLLLTLPPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQI
KGLTGASGKVALLELGCGTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVV
APGEDMRQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFFWEHVAEPYGSWAFM
WQQVFEPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFP
SSKALICSFPSLQLEQATHQPIYLPRLGT

107/310
FIGURE 104

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGAGTTGCCTCATCGCAGG
CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCCTCGCCAGCTTCTGTAGATAAGGGTTAAAA
ACTAATATTTATATGACAGAAGAAAAAGATGTCATTCCGTAAAGTAAACATCATCATCTTGG
TCCTGGCTGTTGCTCTCTTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG
TTAAGGAATGAGGTTACAGATTCAGGAATTGTAGGGCCTCAACCTATAGACTTTGTCCCAA
TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCCCTGTGGTCATCGCTGCATCTG
AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTGAGCACAACACTCGCTCCAAT
GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCTGGCTCAACAG
TGATTCCCTGAAAAGCATCAGATACAAAATTGTCAATTTTGACCCTAAACTTTTGGAAGGAA
AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTAC
TTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCA
AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG
AAGATTGTGATTGAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACCAGTACAATTAC
ATTGGCTATCTTGACTATAAAAAGGAAAGAATTGTAAGCTTTCCATGAAAGCCAGCACTTG
CTCATTTAATCCTGGAGTTTTTTGTTGCAAACCTGACGGAATGGAAACGACAGAATATAACTA
ACCAACTGGAAAAATGGATGAAACTCAATGTAGAAGAGGGACTGTATAGCAGAACCCTGGCT
GGTAGCATCACAAACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC
TATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTA
AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT
ACTGATGTTTGGGAAAAATGGTATATTCCAGACCCAACAGGCAAATTCAACCTAATCCGAAG
ATATACCGAGATCTCAACATAAAGTGAACAGAAATTTGAACTGTAAGCAAGCATTCTCAG
GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA
GCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAACTGCCCTGTCTGGCAGTC
AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTTTCTT
ACTACAATGCTGAATGACTGGAAAGAAGAACTGATATGGCTAGTTCAGCTAGCTGGTACAGA
TAATTCAAACTGCTGTTGGTTTTAATTTTGTAACCTGTGGCCTGATCTGTAAATAAACTT
ACATTTTTTC

—108/310

FIGURE 105

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTDSGIVGPQPIDFVPNALRHAVDGR
QEEIPVVIAASEDRLGGAIAAINS IQHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK
IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNT
ALKPGHAAAFSEDCDSASTKV VIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA
NLTEWKRQNI TNQLEKWMKLNVEEGLYSRTL AGSITTPPLLIVFYQQHSTIDPMWNVRLGS
SAGKRYSPQFVKA AKLLHWNGHLKPWGRTASYTDVWEKWI PDPTGKFNLIRRYTEISNIK

109/310

FIGURE 106

TGGTTTTTGCCCCATAAATTCCCTCAGCTTGAGCAGTTTGTTAAGGAATGAGGTTACAGATT
CAGGAATTNTAGGNCCTCAACCTNTAGANTTTGTCCCAAATGTTCTCCGACATGCAGTAGAT
GGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATNTGAAGACAGGCTTGGGGGGGCCAT
TGCAGCTATAAACAGCATTTCAGCACAACACTCGNTCCAATGTGATTTTCTACATTGTTACTC
TCAACAATACAGCAGACCATNTCCGGTCTGGNTCAACAGTGATTCCCTGAAAAGCATCAGA
TACAAAATTGTCAATTTTGACCCTAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCA
GGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGTTCCCAGCG
CAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTAC
AATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTAGCCTCTAC
TAAAGTTGTCATCCGTGGAGCAGGAAA

110 / 310

FIGURE 107

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCGCGGCTGCCACG
GAGCTAGAGGGCAAGTGTGCTCGGCCCAGCGTGCAGGGAACGCGGGCGGCCAGACAACGGGC
TGGGCTCCGGGGCCTGCGGCGCGGGCGCTGAGCTGGCAGGGCGGGTCCGGGCGCGGGCTGCA
TCCGCATCTCCTCCATCGCCTGCAGTAAGGGCGGCCGCGGCGAGCCTTTGAGGGGAACGACT
TGTCGGAGCCCTAACCCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGGAGCGTCACATCACT
TTCCGATCACTTCAAAGTGGTTAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT
CCGTAAAGTAAACATCATCATCTTGGTCCTGGGCTGTTGCTCTCTTCTTACTGGTTTTGCAC
CATAACTTCCTCAGCTTGAGGCAGTTTGTTAAGGAATGAGGTTACAGATTGAGGAATTGTAG
GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA
GGAGATTCCCTGTGGTCATCGCTGCATCTGAAGACAGGCTTGGGGGGGCCATTGCAGCTATAA
ACAGCATTGAGCACAACACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA
GCAGACCATCTCCGGTCCCTGGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTG
TCAATTTTGACCCTAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC
ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCCAGCGCAAAGAAGG
CCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA
CTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTGAGCCTCTACTAAAGTTGT
CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA
TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTGCAAAC
CTGACGGAATGGAAACGACAGAATATAACTAACCACTGGAAAAATGGATGAAACTCAATGT
AGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAAACACCTCCTCTGCTTATCG
TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT
GCTGGAAAACGATATTCACCTCAGTTTGTAAGGCTGCCAAGTTACTCCATTGGAATGGACA
TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGGAAAAATGGTATATTCCA
GACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAAA
CAGAATTTGAACTGTAAGCAAGCATTTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGGAAG
TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG
GTAAAGATGACAACTGCCCTGTCTGGCAGTCAGCTTCCCAGACAGACTATAGACTATAAAT
ATGTCTCCATCTGCCTTACCAAGTGTTTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA
CTGATATGGCTAGTTCAGCTAGCTGGTACAGATAATTCAAACTGCTGTTGGTTTTAATTTT
GTAACCTGTGGCCTGATCTGTAAATAAACTTACATTTTTCAATAGGTAAAAA

111/310

FIGURE 108

CTGCAGGTAGACATCTCCACTGCCCAGGAATCACTGAGCGTGCAGACAGCACAGCCTCCTCT
GAAGGCCGGCCATACCAGAGTCCTGCCTCGGCATGGGCCTCACCATTGAGGCAGCTCCACTG
TCTGTGCTGGTCTGAGGGTGCTGCCTGTCAATGGGGGCAGCCATCTCCCAGGGGGCCCTCATC
GCCATCGTCTGCAACGGTCTCGTGGGCTTCTTGCTGCTGCTGCTCTGGGTATCCTCTGCTG
GGCCTGCCATTCTCGTCTGCCGACGTTGACTCTCTCTCTGAATCCAGTCCCAACTCCAGCCC
TGGCCCCTGTCCTGAGAAGGCCCCACCACCCCAGAAGCCCAGCCATGAAGGCAGCTACCTGC
TGCAGCCCTGAAGGCCCCCTGGCCTAGCCTGGAGCCCAGGACCTAAGTCCACCTCACCTAGAG
CCTGGAATTAGGATCCCAGAGTTCAGCCAGCCTGGGGTCCAGAACTCAAGAGTCCGCCTGCT
TGGAGCTGGACCCAGCGGCCCCAGAGTCTAGCCAGCTTGGCTCCAATAGGAGCTCAGTGGCCC
TAAGGAGATGGGCCTGGGGTGGGGGCTTATGAGTTGGTGCTAGAGCCAGGGCCATCTGGACT
ATGCTCCATCCCAAGGGCCAAGGGTCAGGGGCCGGGTCCACTCTTTCCCTAGGCTGAGCACC
TCTAGGCCCTCTAGGTTGGGGAAGCAAACCTGGAACCCATGGCAATAATAGGAGGGTGTCCAG
GCTGGGCCCCCTCCCCTGGTCCTCCAGTGTTGCTGGATAATAAATGGAACCTATGGCTCTAA
AAAAAAAAAAAAAAAAAAAA

112 / 310

FIGURE 109

MGA AISQ GALIA IVCNGLVGFL LLLLVILCWACHSRLPTLTLSLNPVPTPALAPVLRPHH

PRSPAMKAATCCSPEGPWPSLEPRT

113/310

FIGURE 110

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCA
GTTCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTA
CTCCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGA
TCATGTCCGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCC
ATGATGTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTATTGGTTAT
TTTGGGATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACC
TCAGCATAGAATTGGACACAGAAAGGGAAAATATGAAGTGCGTGCTGGGGTTTGCTATCGTA
TCCACAGGCATCACGGCAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATT
GACAGTTGAGCTTTTCCAAATCACAAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCC
AGCCACTGTGGACATTTGCCATCCTCATTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTG
AGCCTGGGAACTGCAGGAGCTGCCCAGGTTATGGAAGGCGGCCAAGTGGAATATAAGCCCCCT
TTCGGGCATTCCGTACATGTGGTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCA
TCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAGTGGTTACTTGTTATTTCAACAGAAGT
AAAAATGATCCTCCTGATCATCCCATCCTTTCTGTCTCTCTCCATTCTCTTCTTCTACCATCA
AGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAGGATTCCGAGAATCATTGTCA
TGACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGA
TGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATA
TACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAGATGCATTCAAATCT
TGTCCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTGGAGACTTCATAATTTTTCTA
GGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTACAATCG
GGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTTGCCTACTTAGTAGCCC
ATAGTTTTTTATCTGTGTTTGAACTGTGCTGGATGCACTTTTCCTGTGTTTTGCTGTTGAT
CTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTT
CGTAAAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGA
ATGAGGAGGGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCT
GGAAACATTTCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTT
AGTGAATTTTTTTTTTAAAGACCTAATAAACCTATTCTTCCTCAAAA

114/310

FIGURE 111

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVILGLLFVCGVLWWLYDYTNDL
SIELDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQ
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI
LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVM
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKIL
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN
EEGTELQAIVR

115/310

FIGURE 113

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKC
PAGCQDPKYHVYGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSL
SLPRWRESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQ
LLAVTVAVATPTTLPRPSPSAASTTIPRPQSVGHRSEQMDLWSTATYTSSQNRPRADPGIQ
RQDPGAAAFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFR
IQKQLLADVAQALDIGPAGPLMGVVQYGDNPATFNLKTHHTNSRDLKTAIEKITQRGGLSNV
GRAISFVTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAE
NEKQYVVEPNFANKAVCRTNGFYSLHVQSWFGLHKTQLPLVKRVCDTDRLACSKTCLNSADI
GFVIDGSSSVGTGNFRTVLQFVTNLTKEFEISDTRIGAVQYTYEQRLEFGFDKYSSKPD
LNAIKRVGYWSGGTSTGAAINFALEQLFKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKG
VITYAIGVAWAAQEELEVIATHPARDHSSFFVDEFNLHQYVPRIIQNICTEFNSQPRN

117/310

FIGURE 114

CAGGATGAACTGGTTGCAGTGGCTGCTGCTGCTGCGGGGGCGCTGAGAGGACACGAGCTCTA
TGCCTTTCCGGCTGCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGT
GCGCCAGGTCCCGACGGCTCCGCGCCAGATCCCGCCCACTACAGTTTTTCTCTGACTCTAAT
TGATGCACTGGACACCTTGCTGATTTTGGGGAATGTCTCAGAATTCCAAAGAGTGGTTGAAG
TGCTCCAGGACAGCGTGGACTTTGATATTGATGTGAACGCCTCTGTGTTTGAACAAACATT
CGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGA
GGCTGGATGGCCCTGTTCCGGGCCTCTCCTGAGAATGGCTGAGGAGGCGGCCCGAAACTCC
TCCCAGCCTTTCAGACCCCCACTGGCATGCCATATGGAACAGTGAACTTACTTCATGGCGTG
AACCCAGGAGAGACCCCTGTCACCTGTACGGCAGGGATTGGGACCTTCATTGTTGAATTTGC
CACCTTGAGCAGCCTCACTGGTGACCCGGTGTTTGAAGATGTGGCCAGAGTGGCTTTGATGC
GCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTTCGGCAACCACATTGATGTGCTCACTGGC
AAGTGGGTGGCCAGGACGCAGGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTGGT
GAAAGGAGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCTAGAGTATAACAAAG
CCATCCGGAACACACCCGCTTCGATGACTGGTACCTGTGGGTTCAGATGTACAAGGGGACT
GTGTCCATGCCAGTCTTCCAGTCCTTGGAGGCCTACTGGCCTGGTCTTCAGAGCCTCATTGG
AGACATTGACAATGCCATGAGGACCTTCCTCACTACTACACTGSTATGGAAGCAGTTTGGGG
GGCTCCCGGAATTCTACAACATTCCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTACCCA
CTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCACCT
CCTAGAACTCGGAAGAGATGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGC GGAT
TTGCAACAATCAAAGATCTGCGAGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCCTG
GCCGAGACTGTGAAATACCTCTACCTCCTGTTTGACCCAACCAACTTCATCCACAACAATGG
GTCCACCTTCGACGCGGTGATCACCCCTATGGGGAGTGCATCCTGGGGGCTGGGGGGTACA
TCTTCAACACAGAAGCTCACCCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAAG
GAAGAGCAGTGGGAGGTGGAGGACTTGATGAGGGAATTCTACTCTCTCAAACGGAGCAGGTC
GAAATTTTCAGAAAAACACTGTTAGTTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACAC
TCTTCTCACCAGAAAACCATGACCAGGCAAGGGAGAGGAAGCCTGCCAAACAGAAGGTCCCA
CTTCTCAGCTGCCCCAGTCAGCCCTTCACCTCCAAGTTGGCATTACTGGGACAGGTTTTCTCT
AGACTCCTCATATAACCACTGGATAATTTTTTTATTTTTATTTTTTTGAGGCTAAACTATAATA
AATTGCTTTTGGCTATCATAAAA

118 / 310

FIGURE 115

MPFRLLIPLGLLCALLPQHHGAPGPDGSAPDPAHYSFSLTLIDALDTLLILGNVSEFQRVVE
VLQDSVDFDIDVNASVFETNIRVVGLLSAHLLSKKAGVEVEAGWPCSGPLLRMAEEAARKL
LPAFQTPTGMPYGTVNLLHGVNPGETPVTCTAGIGTFIVEFATLSSLTGDPVFEDVARVALM
RLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFEYLVKGAILLQDKKLMAMFLEYNK
AIRNYTRFDDWYLVWQMYKGTVSMPVFSLEAYWPGLSLIGDIDNAMRTFLNYTIVWKQFG
GLPEFYNIPOGYTVEKREGYPLRPELIESAMYLYRATGDPTLLELGRDAVESIEKISKVECG
FATIKDLRDHKLDNRMESFFLAETVKYLYLLFDPTNFIHNNGSTFDAVITPYGECILGAGGY
IFNTEAHPIDLAALHCCQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPGT
LFSPENHDQARERKPAKQKVPLLSCPSQPFTSKLALLGQVFLDSS

149 / 310
FIGURE 116

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
AACAGTGTACTATTCTGTCTGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATAACAACCTTCGTGTGAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGAGTGG
GGGTATTCCAGTGACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCTTTGTTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTCTGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG
TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCAGAAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCTGAGGAACTCCTCAGGGCCTGGAT
CTCATAGGTTTGGCGAAGGGCCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACACAGAGTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAACTGGTGACACTCTACAGTCTGACTGATTGAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGCCCACTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACTTCAAGCCCAATGCCGGTGAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTCACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACTACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTC
TGTTGGTAAAGTACAGAATTGAGCAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
AAAAAAAAA

-120/310

FIGURE 117

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHELLMWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS AW
SILKHPFNRNSTILTRPGMEITKDG FHLVIELEDLGPQFEFLVAYWRREP GAEEHV KMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLV LALFAFVG FMLILV
VVPLFVWKMG RLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation sites.

amino acids 40-43 and 134-137

Tissue factor proteins homology.

amino acids 92-119

Integrins alpha chain protein homology.

amino acids 232-262

121/310

FIGURE 118

TCCTGCTGATGCACATCTGGGTTTGGCAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTGAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAAGTGAAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTGGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCCAC
ATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

-122/310

FIGURE 119

CGGACGCGTG GGGCCGCCACCTCCGGAACAAGCCATGGTGCGGCGACGGTGGCAGCGGCGTG
GCTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGG
TCAACATCCGGGGCAAACCTGGTGTGCTGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTG
AATGTGGCCAGCGAGTGCGGCTTACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCG
AGACCTGGGCCCCCACCACCTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTTGGCCAACAGG
AGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCCCGGCACCTACAGTGTCTCATTCCCC
ATGTTTAGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCA
GACTTCTGGGAAGGAGCCCACCTGGAACCTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGG
TGGTAGGGGCTTGGGACCCAACCTGTGTCAGTGGAGGAGGTCAGACCCCAGATCACAGCGCTC
GTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTTATTAACCACCGCGTCTCCTCCTCCACCA
CCTCATCCCGCCACCTGTGTGGGGCTGACCAATGCAAACCTCAAATGGTGCTTCAAAGGGAG
AGACCCACTGACTCTCCTTCCTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAA
AAATTCTAGTATTTTGATTATTTGAATCTTACAGCAACAAATAGGAACTCCTGGCCAATGAG
AGCTCTTGACCAGTGAATCACCAGCCGATACGAACGTCTTGCCAACAAAAATGTGTGGCAAA
TAGAAGTATATCAAGCAATAATCTCCACCCAAGGCTTCTGTAACTGGGACCAATGATTAC
CTCATAGGGCTGTTGTGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGC
CAAATAGGAGGCATTCAATGAACATTTTTTGCATATAAACCAAAAAATAACTTGTTATCAAT
AAAAACTTGCATCCAACATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTT
GTTATTTCTCTGTATTATTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCA
AACAAACCTCACGATATAAAATAAAAATGAAAGTATCCTCCTCAAAAA

123 / 310

FIGURE 120

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVNVASECGFTDQ
HYRALQQLQORDLGPHHFNVLAFFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTG
AHPAFKYLAQTSGKEPTWNFWKYLVPDGGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

124/310

FIGURE 121

CGGACGCGTGCGGGCGGGCCGGGACGCAGGGCAAAGCGAGCCATGGCTGTCTACGTCGGGATGC
TGCGCCTGGGGAGGCTGTGCGCCGGGAGCTCGGGGGTGCTGGGGGCCCGGGCCGCCCTCTCT
CGGAGTTGGCAGGAAGCCAGGTTGCAGGGTGTCCGCTTCCTCAGTTCCAGAGAGGTGGATCG
CATGGTCTCCACGCCCATCGGAGGCCTCAGCTACGTTCAAGGGTGCACCAAAAAGCATCTTA
ACAGCAAGACTGTGGGCCAGTGCCTGGAGACCACAGCACAGAGGGTCCCAGAACGAGAGGCC
TTGGTCGTCTCCATGAAGACGTCAGGTTGACCTTTGCCCAACTCAAGGAGGAGGTGGACAA
AGCTGCTTCTGGCCTCCTGAGCATTGGCCTCTGCAAAGGTGACCGGCTGGGCATGTGGGGAC
CTAACTCCTATGCATGGGTGCTCATGCAGTTGGCCACCGCCCAGGCGGGCATCATTCTGGTG
TCTGTGAACCCAGCCTACCAGGCTATGGAAGTGGAGTATGTCCTCAAGAAGGTGGGCTGCAA
GGCCCTTGTTGTTCCCCAAGCAATTCAAGACCCAGCAATACTACAACGTCCTGAAGCAGATCT
GTCCAGAAGTGGAGAATGCCCAGCCAGGGGCCCTGAAGAGTCAGAGGCTCCAGATCTGACC
ACAGTCATCTCGGTGGATGCCCCCTTTGCCGGGGACCCCTGCTCCTGGATGAAGTGGTGGCGGC
TGGCAGCACACGGCAGCATCTGGACCAGCTCCAATACAACCAGCAGTTCTGTCTGCGCATG
ACCCCATCAACATCCAGTTACCTCGGGGACAACAGGCAGCCCCAAGGGGGCCACCCTCTCC
CACTACAACATTGTCAACAACCTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGAC
ACCAGAGCAGTTGCGGATGATCCTGCCAACCCCTGTACCATTGCCTGGGTTCGTTGGCAG
GCACAATGATGTGTCTGATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGC
AAGAAGGCACTGGAGGCCATCAGCAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGAT
GTTCTGTTGACATTCTGAACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAG
GTGTCAATTGCTGGGTCCCCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAAT
ATGAAGGACCTGGTGGTTGCTTATGGAACACAGAGAACAGTCCCGTGACATTCGCGCACTT
CCCTGAGGACACTGTGGAGCAGAAGGCAGAAAGCGTGGGCAGAATTATGCCTCACACGGAGG
CCCGGATCATGAACATGGAGGCAGGGACGCTGGCAAAGCTGAACACGCCCCGGGGAGCTGTGC
ATCCGAGGGTACTGCGTCATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGT
GGATCAGGACAAGTGGTATTGGACAGGAGATGTGCCACAATGAATGAGCAGGGCTTCTGCA
AGATCGTGGGCCGCTCTAAGGATATGATCATCCGGGGTGGTGAGAACATCTACCCCGCAGAG
CTCGAGGACTTCTTTCACACACACCCGAAGGTGCAGGAAGTGCAGGTGGTGGGAGTGAAGGA
CGATCGGATGGGGGAAGAGATTTGTGCCTGCATTCCGGCTGAAGGACGGGGAGGAGACCACGG
TGGAGGAGATAAAAGCTTTCTGCAAAGGGAAGATCTCTCACTTCAAGATTCCGAAGTACATC
GTGTTTGTCAAACTACCCCTCACCATTTCAGGAAAGATCCAGAAATTCAACTTCGAGA
GCAGATGGAACGACATCTAAATCTGTGAATAAAGCAGCAGGCCTGTCCTGGCCGGTTGGCTT
GACTCTCTCCTGTCAGAATGCAACCTGGCTTTATGCACCTAGATGTCCCCAGCACCCAGTTC
TGAGCCAGGCACATCAAATGTCAAGGAATTGACTGAACGAACTAAGAGCTCCTGGATGGGTC
CGGGAACCTCGCCTGGGCACAAGGTGCCAAAAGGCAGGCAGCCTGCCAGGCCCTCCCTCCTG
TCCATCCCCACATTCCCCTGTCTGTCTTGTGATTTGGCATAAAGAGCTTCTGTTTTCTTT
GAAAAAAAAAAAAAAAAA

125/310

FIGURE 122

MAVYVGMRLRLGRLCAGSSGVLGARAALSRWQEARLQGVRFLLSSREVDRMVSTPIGGLSYVQ
GCTKKHLNSKTVGQCLETTAQRVPEREALVVLHEDVRLTFAQLKEEVDKAASGLLSIGLCKG
DRLGMWGPNSYAWVLMQLATAQAGIILVSVNPAYQAMELEYVLKKVGCKALVFPKQFKTQQY
YNVLKQICPEVENAQPGALKSQRLPDLTTVISVDAPLPGTLLLDEVVAAGSTRQHLDQLQYN
QQFLSCHDPINIQFTSGTTGSPKGATLSHYNIVNNSNILGERLKLHEKTPEQLRMILPNPLY
HCLGSVAGTMMCLMYGATLILASPIFNGKKALEAISRERGTFLYGTPTMFVDILNQPDFSSY
DISTMCGGVIAGSPAPPELIRAIINKINMKDLVVAYGTTENSPVTFAHFPEDTVEQKAESVG
RIMPHTEARIMNMEAGTLAKLNTPGELCIRGYCVMLGYWGEPQKTEEAVDQDKWYWTGDVAT
MNEQGFCKIVGRSKDMIIRGGENIYPAELEDDFFHHPKVQEVQVVGKDDRMGEEICACIRL
KDGEETTVEEIKAFCKGKISHFKIPKYIVFVTNYPLTISGKIQKFLREQMERHLNL

126/310

FIGURE 123

CAACTCCAACATTTTAGGAGAGCGCCTGAACTGCATGAGAAGACACCAGAGCAGTTGCGGA
TGATCCTGCCCCAACCCCTGTACCATTGCCTGGGTTCCGTGGCAGGCACAATGATGTGTCTG
ATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGCAAGAAGGCACTGGAGGC
CATCAGCAGAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGATGTTCGTGGACATTCTGA
ACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAGGTGTCATTGCTGGGTCC
CCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAATATGAAGGACCTGGTGGT
TGCTTATGGAACCACAGAGAACAGTCCCGTGACATTCGCGCACTTCCCTGAGGACACTGTGG
AGCAGAAGGCAGAAAGCGTGCGCAGAATTATGCCTCACACGGAGGCGCGGATCATGAACATG
GAGGCAGGGACGCTGGCAAAGCTGAACACGCCCCGGGGAGCTGTGCATCCGAGGGTACTGCGT
CATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGTGGATCAGGACAAGTGGT
ATTGGACAGGAGATGTGCCAC

127/310

FIGURE 124

GAGCAGGACGGAGCCATGGACCCCGCCAGGAAAGCAGGTGCCAGGCCATGATCTGGACTGC
AGGCTGGCTGCTGCTGCTGCTGCTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCG
TGCAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGC
GTGGACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGC
AGTGCGGGGTTGCGGTTGCGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGC
TTCTGGCGTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTC
ACCTCGCGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCCAACGGCGTGGAGTG
CTACAGCTGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTCGTGAGCT
GCTACAACGCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTCACTTGACGGCA
GCTAATGTGACTGTGTCTTGCTGCTCGGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGA
TGGAGTAACAGGCCCAGGGTTCACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCGCTGTAACT
CTGACCTCCGCAACAAGACCTACTTCTCCCCTCGAATCCCACCCCTTGTCGGGCTGCCCCCT
CCAGAGCCCACGACTGTGGCCTCAACCACATCTGTCACCACTTCTACCTCGGCCCCAGTGAG
ACCCACATCCACCACCAAAACCCATGCCAGCGCCAACAGTCAGACTCCGAGACAGGGAGTAG
AACACGAGGCCTCCCGGGATGAGGAGCCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGAC
CGCAGCAATTCAGGGCAGTATCCTGCAAAAGGGGGGCCCCAGCAGCCCCATAATAAAGGCTG
TGTGGCTCCCACAGCTGGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCTACTGT
GAGCTTCTCCACCTGGAAATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCT
CATCACTTCCTGTTCCCACCACTGGACTGGGCTGGCCCAGCCCCTGTTTTTCCAACATTCCC
CAGTATCCCCAGCTTCTGCTGCGCTGGTTTGCGGCTTTGGGAAATAAAATACCGTTGTATAT
ATTCTGCCAGGGGTGTTCTAGCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGCTC
TCCGCTTGTCCTCTGTGATGTTAGGACAGAGTGAGAGAAGTCAGCTGTACGGGGAAGGTG
AGAGAGAGGATGCTAAGCTTCCTACTCACTTTCTCCTAGCCAGCCTGGACTTTGGAGCGTG
GGTGGGTGGGACAATGGCTCCCCACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGG
GAATCGGTTCCCCATATGTCTTCCTTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTAC
CCAATTCGCCCTATAGTGAGTCGTA

128/310

FIGURE 125

MDPARKAGAQAMIWTAGWLLLLLLLRGGAQALECYSCVQKADDGCSPNKMKTVKCAPGVDVCT
EAVGAVETIHGQFSLAVRGCGSGLPGKNDRGLDLHGLLAFIQLQQCAQDRCNAKLNLTSRAL
DPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGCFDGNVTLTAANVTV
SLPVRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRIPLVRLPPPEPTT
VASTTSVTTSTSAVRPTSTTKPMPAPTSQTPRQGVHEASRDEEPRLTGGAAGHQDRSNSG
QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

130 / 310

FIGURE 127

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRCK

CSFNQKPRAPGDDEEAQVENLITANATEPQKQRTQEVQPSGGSLWNLRRLLLEPLDANVDA

131/310

FIGURE 128

AAACTTGACGCCATGAAGATCCCGGTCCTTCCTGCCGTGGTGCTCCTCTCCCTCCTGGTGCT
CCTACTCTGCCCAGGGAGCCACCCTGGGTGGTCCTGAGGAAGAAAGCACCATTGAGAATTATG
CGTCACGACCCGAGGCCTTTAACACCCCGTTTCCTGAACATCGACAAATTGCGATCTGCGTTT
AAGGCTGATGAGTTCCTGAACTGGCACGCCCTCTTTGAGTCTATCAAAGGAACTTCCTTT
CCTCAACTGGGATGCCTTTCCTAAGCTGAAAGGACTGAGGAGCGCAACTCCTGATGCCCAGT
GACCATGACCTCCACTGGAAGAGGGGGCTAGCGTGAGCGCTGATTCTCAACCTACCATAACT
CTTTCCTGCCTCAGGAACTCCAATAAAACATTTTCATCCAAA

132/310

FIGURE 129

MKIPVLPVVLLSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSFAKDE
FLNWHALFESI KRKLPFLNWDAPFKLKGLRSATPDAQ

133/310

FIGURE 130

CAGTTCTGAAATCAATGGAGTTAATTTAGGGAATACAAACCAGCCATGGGGGGTGGAGATTGC
CTTTGCCTCAGTGATTCTCACCTGCCTCTCCCTTCTGGCAGCAGGAGTCTCCCAGGTTGTTC
TTCTCCAGCCAGTTCCAACCTCAGGAGACAGGTCCCAAGGCCATGGGAGATCTCTCCTGTGGC
TTTGCCGGCCACTCATGAGAGTGTTTTTGTGTAAAGTATTTTTTTAGAACTGTTGACTTCT
TCATGATTTAATAACCATCCTTTGCGAAGTTTTATGAGGCTTTAGGGGAATGTCAACCCTCA
AATTTTTGTTATACTAGATGGCTTCCATTTACCCACCACTATTTTAAGGTCCCTTTATTTTT
AGGTTCAAGGTCATTTGACTTGAGAAAGTGCCCTTCTGCAGCTTCATTGATTTTGTATTATC
TTCATATTAATTGTAACGATTAAAAAGAATAAGAGCACGCAGACCTCTAGGAGAAATATTT
TATCCCTGGGTGCCCCCTGACACATTTATGTAGTGATCCCACAAATGTGATTGTTAATTTAAA
TGTTATTCTAATATTAGTACATTCAGTTGTGATGTAATATGAATAACCAGAATCTATTTCTT
AAAAGTTTTGAGTATATTTTTTCAACTAGATATTTGTATAGAAAGACTGAATAGTGATG

134/310

FIGURE 131

MGVEIAFASVILTCLSLAAGVSQVLLQPVPTQETGPKAMGDLSCGFAGHS

135/310

FIGURE 132

GGGGAATCTGCAGTAGGTCTGCCGGCGATGGAGTGGTGGGCTAGCTCGCCGCTTCGGCTCTG
GCTGCTGTTGTTCTCCTGCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAAATGGA
AAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGTCAAAAC
TGCAGCTGCTACCATGGTGT CATAGAAGAGGATCTAACTCCTTTCCGAGGAGGCATCTCCAG
GAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCACTAAGAACA
GACTGTACCGGGAAAAATGACTGCATGTTCCCCTCAAGGTGTAGTGGTGTGAGCACTTTATT
TTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTACGAGATTATCCTCA
GGTTCCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACC
ATGATATCATGTATCCTGCTTGGACATTTTGGGAAGGGGGACCTGCTGTTTGGCCAATTTAT
CCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGGTAAGGTCAGCAGCACAGTG
GCCATGGAAAAAGAAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAG
ATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAAACCTTGTTGATGCAGAATACACCAAAAAC
CAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGT
GGATCACTGCAAATACAAGTATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTA
AACACCTCTTCCTGTGTGGCTCACTTGTTTTCCATGTTGGTGATGAGTGGCTAGAATTCTTC
TATCCACAGCTGAAGCCATGGGTTCACTATATCCCAGTCAAAACAGATCTCTCCAATGTCCA
AGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAA
GCCAGTTTATTAGGAACCATTGTCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTG
AGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTAT
TCCCAAAATGTTGAAAACCTGAACTATAGTAGTCATCATAGGACCATAGTCCTCTTTGTGGCA
ACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATA
TCTGCTATCAAGCCAAATACCTGGTTTTCTTATCATGCTGCACCCAGAGCAACTCTTGAGA
AAGATTTAAAATGTGTCTAATACACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGGA
CCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTCATTTTCTTAAGACCAATC
ACAGCTTGTGCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCA
TGTGATGATGCCCTTTGTCCCATTTTGGAGCAGAAAATTCGTCATTTGGAAGTAGTACAA
CTCATTGCTGGAATTGTGAAATTATTCAAGGCGTGATCTCTGTCACTTTATTTTAATGTAGG
AAACCCTATGGGGTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGG
TAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTGTAAAACCATAAACTCTGTTACTCAG
GAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT
TCAGGTTCCCTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

136/310

FIGURE 133

MEWWASSPLRLWLLLFLLP
SAQGRQKESGSKWKVFIDQ
INRSLENYEPCSSQNCSCYH
GVIE
EDLTPFRGGISRKMAEVRR
KLGTHYQITKNRLYREND
CMFPSRCSGVEHFILEVIG
RLPD
MEMVINVRDYPQVPKWME
PAIPVFSFSKTSEYHDIM
YPAWTFWEGGPAVWPIYPT
GLGRWDL
FREDLVRSAQWPWKKNST
AYFRGSRTSPERDPLILL
SRKNPKLVDAEYTKNQAW
KSMKDT
LGKPAAKDVHLVDHCKYK
YLFNFRGVAASFRFKHLF
CGSLVFHVGDEWLEFFYP
QLKPWVH
YIPVKTDLSNVQELLQFV
KANDDVAQEIAERGSQFI
RNHLQMDDITCYWENLLS
EYSKFLSY
NVTRRKGYDQIIPKMLK
TEL

-137/310

FIGURE 134

CACCCCTCCATTTCTCGCCATGGCCCCCTGCACTGCTCCTGATCCCTGCTGCCCTCGCCTCTT
TCATCCTGGCCTTTGGCACCGGAGTGGAGTTCGTGCGCTTTACCTCCCTTCGGCCACTTCTT
GGAGGGATCCCGGAGTCTGGTGGTCCGGATGCCCGCCAGGGATGGCTGGCTGCCCTGCAGGA
CCGCAGCATCCTTGCCCCCCTGGCATGGGATCTGGGGCTCCTGCTTCTATTTGTTGGGCAGC
ACAGCCTCATGGCAGCTGAAAGAGTGAAGGCATGGACATCCCGGTACTTTGGGGTCCCTCAG
AGGTCACTGTATGTGGCCTGCACTGCCCTGGCCTTGCACTGGTGATGCGGTACTGGGAGCC
CATACCCAAAGGCCCTGTGTTGTGGGAGGCTCGGGCTGAGCCATGGGCCACCTGGGTGCCGC
TCCTCTGCTTTGTGCTCCATGTCATCTCCTGGCTCCTCATCTTTAGCATCCTTCTCGTCTTT
GACTATGCTGAGCTCATGGGCCTCAAACAGGTATACTACCATGTGCTGGGGCTGGGCGAGCC
TCTGGCCCTGAAGTCTCCCCGGGCTCTCAGACTCTTCTCCACCTGCGCCACCCAGTGTGTG
TGGAGCTGCTGACAGTGCTGTGGGTGGTGCCTACCCTGGGCACGGACCGTCTCCTCCTTGCT
TTCCTCCTTACCCTCTACCTGGGCCTGGCTCACGGGCTTGATCAGCAAGACCTCCGCTACCT
CCGGGCCCAGCTACAAAGAAAATCCACCTGCTCTCTCGGCCCCAGGATGGGGAGGCAGAGT
GAGGAGCTCACTCTGGTTACAAGCCCTGTTCTTCTCTCCCACTGAATTCTAAATCCTTAAC
ATCCAGGCCCTGGCTGCTTCATGCCAGAGGCCCAAATCCATGGACTGAAGGAGATGCCCCCTT
CTACTACTTGAGACTTTATTCTCTGGGTCCAGCTCCATACCCTAAATTCTGAGTTTCAGCCA
CTGAACTCCAAGGTCCACTTCTCACCAGCAAGGAAGAGTGGGGTATGGAAGTCATCTGTCCC
TTCCTGTTTAGAGCATGACACTCTCCCCCTCAACAGCCTCCTGAGAAGGAAAGGATCTGCC
CTGACCACTCCCCTGGCACTGTTACTTGCCTCTGCGCCTCAGGGGTCCCCTTCTGCACCGCT
GGCTTCCACTCCAAGAAGGTGGACCAGGGTCTGCAAGTTCAACGGTCATAGCTGTCCCTCCA
GGCCCCAACCTTGCCTCACCCTCCCGGCCCTAGTCTCTGCACCTCCTTAGGCCCTGCCTCT
GGGCTCAGACCCCAACCTAGTCAAGGGGATTCTCCTGCTCTTAACTCGATGACTTGGGGCTC
CCTGCTCTCCCGAGGAAGATGCTCTGCAGGAAAATAAAAGTCAGCCTTTTTCTAAAAAAA

-138/310

FIGURE 135

MAPALLLIPAAASFILAFGTGVEFVRFTSLRPLLGGIPESGGPDARQGWLAAALQDRSILAP
LAWDLGLLLLFVVGQHSLMAAERVKAWSRYFGVLQRSLYVACTALALQLVMRYWEPIPKGPV
LWEARAEPWATWVPLLCFVLHVISWLLIFSILLVFDYAELMGLKQVYYHVLGLGEPLALKSP
RALRLFSLRHPVCVELLTVLWVVPTLGTDRLLLAFLLTLYLGLAHGLDQQDLRYLRAQLQR
KLHLLSRPQDGEAE

139/310

FIGURE 136

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGA
AGAAATTGCCAAACCATGTCTTTTTTCTGTTTTTCAGAGTAGTTCAACAACAGATCTGAGTGT
TTTAATTAAGCATGGAATACAGAAAAACAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATT
CCACAGTTTTCTTAGCTCCCTGGACCCGGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTA
TCACGTGGTGCTCTCCGACTACTCACCCGAGTGTAAGAACCTTCGGCTCGCGTGCTTCTG
AGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCTTCCGAGTAGGATGTCACTGAGATCC
CTCAAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCT
TCCCCACTACAATGTGATAGAACGCGTGAACTGGATGTACTTCTATGAGTATGAGCCGATTT
ACAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTT
CTGGTCATTCTGGTGACCTCCACCCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTAC
TTGGGGTGAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAG
AGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCTTAGAGGATGAACACCTTCTTTATGGT
GACATAATCCGACAAGATTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTTATGGC
ATTCAGGTGGGTAACTGAGTTTTGCCCAATGCCAAGTACGTAATGAAGACAGACACTGATG
TTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTT
TTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATAT
TTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTTGGGTTATATAA
TGTCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGTCACGTAAAACCCATCAAGTTT
GAAGATGTTTATGTGCGGATCTGTTTGAATTTATTAAGTGAACATTATATCCAGAAGA
CACAAATCTTTTCTTTCTATATAGAATCCATTGGATGTCTGTCAACTGAGACGTGTGATTG
CAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTTTTGGCAGGTCATGCTAAGGAACACC
ACATGCCATTATTAAGCTTCACATTCTACAAAAAGCCTAGAAGGACAGGATACCTTGTGGAAA
GTGTTAAATAAAGTAGGTAAGTGTGGAAAATTCATGGGGAGGTCAGTGTGCTGGCTTACACTG
AACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTC
AGGCCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAA
GAAATTAATAGGACCAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAAGGG
TGTTACTGAGTTATAAGCTCACTAGGCTGTAAAAACAAAACAATGTAGAGTTTTATTTATTG
AACAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAATTTAAAAATATA
TGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTT
TGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTTACAGTTATTATTATTTAAAATTA
CTTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAGTGAAT
CATTCTTTACATGCAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCAC
TCCATTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTAAAT
ATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAA

140 / 310
FIGURE 137

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQD
FHFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EK
EDKMLALSLEDEHLLYGDII RQDFLD TYNNLT LKTIMAFRWVTEFCPNAKYVMKTD TDVFIN
TGNLVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRD
LVPRIYEMMGHV KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHG
FSSKEIITFWQVMLRNTTCHY

141/310

FIGURE 138

CCTCTGTCCACTGCTTTTCGTGAAGACAAGATGAAGTTCACAATTGTCTTTGCTGGACTTCTT
GGAGTCTTTCTAGCTCCTGCCCTAGCTAACTATAATATCAACGTCAATGATGACAACAACAA
TGCTGGAAGTGGGCAGCAGTCAGTGAGTGTCAACAATGAACACAATGTGGCCAATGTTGACA
ATAACAACGGATGGGACTCCTGGAATTCCATCTGGGATTATGGAAATGGCTTTGCTGCAACC
AGACTCTTTCAAAAGAAGACATGCATTGTGCACAAAATGAACAAGGAAGTCATGCCCTCCAT
TCAATCCCTTGATGCACTGGTCAAGGAAAAGAAGCTTCAGGGTAAGGGACCAGGAGGACCAC
CTCCCAAGGGCCTGATGTACTCAGTCAACCCAAACAAAGTCGATGACCTGAGCAAGTTCGGA
AAAAACATTGCAAACATGTGTTCGTGGGATTCCAACATACATGGCTGAGGAGATGCAAGAGGC
AAGCCTGTTTTTTTACTCAGGAACGTGCTACACGACCAGTGTACTATGGATTGTGGACATTT
CCTTCTGTGGAGACACGGTGGAGAACTTAAACAATTTTTTAAAGCCACTATGGATTTAGTCAT
CTGAATATGCTGTGCAGAAAAAATATGGGCTCCAGTGGTTTTTACCATGTCATTCTGAAATT
TTTCTCTACTAGTTATGTTTGATTTCTTTAAGTTTCAATAAAATCATTTAGCATTGAAAAAA

142/310

FIGURE 139

MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNGWDSWNS
IWDYGNGFAATRLFQKKTCTIVHKMNKEVMPSIQSLDALVKEKKLQGKGPGGPPPKGLMYSVN
PNKVDDLKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN

143/310

FIGURE 140

CATTTCTGAACTAATCGTGTCAGAATTGACTTTGAAAAGCATTGCTTTTTACAGAAGTATA
TTAACTTTTTAGGAGTAATTTCTAGTTTGGATTGTAATATGAAATAATTTAAAAGGGCTTCG
CTCATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTCTTATTGCTTACTGATTTTTT
TGAGTTAAGAGTTGTTATATGCTAGAAATATGAGGATGTGAATATAAATAAGAGAAGAAAAA
GAATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGTGGTTTGTGTACATG
CAAGCTTATAGTTGAAATATTTTTTCAGGAATTACATGAATGACAGTCTTCGAACCAATGTGT
TTGTTTCGATTTCAACCAGAGACTATAGCATGTGCTTGCATCTACCTTGCAGCTAGAGCACTT
CAGATTCCGTTGCCAACTCGTCCCCATTGGTTTCTTCTTTTGGTACTACAGAAGAGGAAAT
CCAGGAAATCTGCATAGAAACACTTAGGCTTTATACCAGAAAAAGCCAACTATGAATTAC
TGGAAAAAGAAGTAGAAAAAGAAAAGTAGCCTTACAAGAAGCCAAATTAAAAGCAAAGGGA
TTGAATCCGGATGGAAGTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCCAGCCTCCAAGCC
ATCATCACCAGAGAAGTAAAAGCTGAAGAGAAATCACCAATCTCCATTAATGTGAAGACAG
TCAAAAAGAACCTGAGGATAGACAACAGGCTTCCAAAAGCCCTTACAATGGTGTAAAGAAA
GACAGCAAGAGAAGTAGAAATAGCAGAAGTGCAAGTCGATCGAGGTCAAGAACACGATCACG
TTCTAGATCACATACTCCAAGAAGACACTATAATAATAGGCGGAGTCGATCTGGAACATACA
GCTCGAGATCAAGAAGCAGGTCCCGCAGTCACAGTGAAAGCCCTCGAAGACATCATAATCAT
GGTTCCTCACCTTAAGGCCAAGCATACCAGAGATGATTTAAAAGTTCAAACAGACATGG
TCATAAAAGGAAAAAATCTCGTTCTCGATCTCAGAGCAAGTCTCGGGATCACTCAGATGCAG
CCAAGAAACACAGGCATGAAAGGGGACATCATAGGGACAGGCGTGAACGATCTCGCTCCTTT
GAGAGGTCCCATAAAAGCAAGCACCATGGTGGCAGTCGCTCAGGACATGGCAGGCACAGGCG
CTGACTTTCTCTTCCTTTGAGCCTGCATCAGTTCTTGGTTTTGCCTATCTACAGTGTGATGT
ATGGACTCAATCAAAAACATTAAACGCAAAGTATTAGGATTTGATTTCTTGAAACCCTCTA
GGTCTCTAGAACACTGAGGACAGTTTCTTTTGAAAAGAACTATGTTAATTTTTTGCACATT
AAAATGCCCTAGCAGTATCTAATTAATAAACCATGGTCAGGTTCAATTGTACTTTATTATAGT
TGTGTATTGTTTATTGCTATAAGAACTGGAGCGTGAATTCTGTAAAATGTATCTTATTTTT
ATACAGATAAAATTGCAGACACTGTTCTATTTAAGTGTTATTTGTTTAAATGATGGTGAAT
ACTTTCCTTAACACTGGTTTGTCTGCATGTGTAAAGATTTTTTACAAGGAAATAAATAACAAT
CTTGTTTTTTCTAAAAAAGT

144/310
FIGURE 141

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLFLFGTTEEEIQEICIETLRLY
TRKKPNYELLEKEVEKRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEK
SPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSTRSRSRSHTPRRHYN
NRRSRSGTYSSRSRSTRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRKKSRSRSQ
SKSRDHSDAAKKHRHERGHHRDRRERSRSFERSHKSKHHGGSRSRSGHGRHRR

145/310
FIGURE 142

TGGGGATAAAGGAAAAATGGTCAGGTATTAATGGCTTAAAGATTATTGGAAGGGGTTTATCA
TTTTTTGAANNTATTCGGGTCANAATTGNCTTTGAAAAGCATTGCTTTTTACAGAAATATAT
TANCTTTTTAGAGTAATTTCTAGTTTGGATTGTAATATGAAATTATTTAAAAGGGCTTCGCT
CATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTNTTATTGCTTACTGATTTTTTTG
AGTTAAGAGTTGTTATATGNTAGAATATGAGGATGTGAATATAAATAAGAGAAGAAAAAGA
ATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGTGGTTTGTTTACATGCA
AGCTTATAGTTGAAATATTTTTCAGGAATTACATGAATGACAGTCTTCGAACCAATGTGTTT
GTTCGATTTCAACCAGAGANTATAGCATGTGCTTGCATCTACCTTGCAGNTAGAGCACTTCA
GATTCCGTTGCCAACTNGTCCCCATTGGTTTCTTCTTTTTGGTACTACAGAAGAGGAAATCC
AGGAAATNTGCATAGAAACACTTAGGCTTTATACCAGAAAAAGCCAACTATGAATTACTG
GAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCNAATTTAAAAGCAAAGGGATT
GAATCCGGATGGAAGTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCC

146/310

FIGURE 143

GGCACGAGGCCTCGTGCCAAGCTTGGCACGAGGGTGCACCGCGTTCTCGCACGCGTCAATGGC
GGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCACCCTCATGCACAGGCTGGCGC
CACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAACGGCAGTTTGTTCGGATACAAGCACCCG
TCTGAGGAGGAGCTTCGGGCCCTGGCGGGGAAGCCGAGGCCAGAGGCAGGAAAGAGCGGTG
GGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCCGAGATGCCCCGTTCCAGCTGG
AGACCTGCCCCCTCACGACCCTGGATGCCCTGGTCTCGCTTCTTCCTGGAGTACCAGTGG
TTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTTCACAGAGGCCTACTACTACAT
GCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGTGCCTGCTCACGGTGACCTTCT
CCATCAAGATGTTCTTGACAGTGACACGGCTGTACTTCAGCGCCGAGGAGGGGGGTGAGCGC
TCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGCCATGCTGGTGCAAGTGGTGCG
GGAGGAGACCCTCGAGCTGGGCCTGGAGCCTGGTCTGGCCAGCATGACCCAGAACTTAGAGC
CACTTCTGAAGAAGCAGGGCTGGGACTGGGCGCTTCCTGTGGCCAAGCTGGCTATCCGCGTG
GGACTGGCAGTGGTGGGCTCTGTGCTGGGTGCCTTCCTCACCTTCCAGGCCTGCGGCTGGC
CCAGACCCACCGGGACGCACTGACCATGTCTGGAGGACAGACCCATGCTGCAGTTCCTCCTGC
ACACCAGCTTCCTGTCTCCCCTGTTTCATCCTGTGGCTCTGGACAAAGCCCATTGCACGGGAC
TTCCTGCACCAGCCGCCGTTTGGGGAGACGCGTTTCTCCCTGCTGTCCGATTCTGCCTTCGA
CTCTGGGCGCCTCTGGTTGCTGGTGGTGCTGTGCCTGCTGCGGCTGGCGGTGACCCGGCCCC
ACCTGCAGGCCTACCTGTGCCTGGCCAAGGCCCGGGTGGAGCAGCTGCGAAGGGAGGCTGGC
CGCATCGAAGCCCGTGAAATCCAGCAGAGGGTGGTCCGAGTCTACTGCTATGTGACCGTGGT
GAGCTTGCAGTACCTGACGCCGCTCATCCTCACCTCAACTGCACACTTCTGCTCAAGACGC
TGGGAGGCTATTCTGGGGCCTGGGCCCAGCTCCTCTACTATCCCCGACCCATCCTCAGCC
AGCGCTGCCCCCATCGGCTCTGGGGAGGACGAAGTCCAGCAGACTGCAGCGCGGATTGCCGG
GGCCCTGGGTGGCCTGCTTACTCCCCTCTTCCTCCGTGGCGTCTGGCCTACCTCATCTGGT
GGACGGCTGCCTGCCAGCTGCTCGCCAGCCTTTTCGGCCTCTACTTCCACCAGCACTTGGCA
GGCTCCTAGCTGCCTGCAGACCCTCCTGGGGCCCTGAGGTCTGTTCTGGGGCAGCGGGACA
CTAGCCTGCCCCCTCTGTTTGGCGCCCCGTGTCCCCAGCTGCAAGGTGGGGCCGGACTCCCC
GGCGTTCCCTTCACCACAGTGCCTGACCCGCGGCCCCCCTTGGACGCCGAGTTTCTGCCTCA
GAACTGTCTCTCCTGGGCCCAGCAGCATGAGGGTCCCGAGGCCATTGTCTCCGAAGCGTATG
TGCCAGGTTTGAGTGGCGAGGGTGATGCTGGCTGCTCTTCTGAACAAATAAAGGAGCATGCC
GATTTTAA

147/310
FIGURE 144

MAVLGVQLVVTLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEEELRALAGKPRPRGRKE
RWANGLSEEKPLSVPRDAPFQLETCPLETTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY
YMLGPAKETNIAVFWCLLTVTFSTIKMFLTVTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV
VREETLELGLPGLASMTQNLPLELLKKQGWDWALPVAKLAIRVGLAVVGSVLGAFLTFPGLR
LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA
FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT
VVSLOYLTPILTLNCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAAARI
AGALGGLLTPLFLRGVLAYLIWWTAAACQLLASLFGLYFHQHLGS

FIGURE 145

CGTTNGCACGCGTCAATGGCGGTCCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCAC
CCTCATGCACAGGCTGGCGCCACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAACGGCAGTT
TGTTCCGATACAAGCACCCGTNTTGAGGAGGAGCTTCGGGGCCCTGGCGGGGAAGCCGAGGCC
CAGAGGCAGGAAAGAGCGGTGGGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCC
GAGATGCCCCGTTCCAGCTGGAGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGC
TTCTTCCTGGAGTACCAGTGGTTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGACCTCTT
CACAGAGGCCTACTACTACATGCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGT
GCCTGCTCACAGTGACCTTCTCCATCAAGATGTTCTGACAGTGACACGGCTGTACTTCAGC
GCCGAGGAGGGGGGTGAGCGCTCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGC
CATGCTGGTGCAAGCG

149/310

FIGURE 146

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTA
TTAACGTGGCTTAATCTGAAGGTTCTCAGTCAAATTCTTTGTGATCTACTGATTGTGGGGGC
ATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGGTTTGGGCCCTTGTAGCTGACAGAAGGTGGC
CAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGCTTCTGTTGCTGGTCTTGCCTTGG
CTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTGTATTGAGAACTCTG
TAAAGGTGCCTCCCCTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTC
CAGACGGCTGTGCGAGCCTCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACC
ATCTCCTTAATGACAGACGAGCCTGGCCTAGACAACCCTGCCTACGTGTCCTCGGCAGAGGA
CGGGCAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAACTAGGGCACGGCCCT
TTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTTCTT
CGAAGGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGGAAAATTCTGA
AAACACCACTGCCCCCTGAAGTCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAAATTA
CCAGCATCAAGATCAATCGAGTAGATCCAGTGAAAGCCTCTCTATTAGGCTGGTGGGAGGT
AGCGAAACCCCACTGGTCCATATCATTATCCAACACATTTATCGTGATGGGGTGATCGCCAG
AGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATG
TCCCTCACAACCTACGCTGTGCGTCTCCTGCGGCAGCCCTGCCAGGTGCTGTGGCTGACTGTG
ATGCGTGAACAGAAGTTCGCGCAGCAGGAACAATGGACAGGCCCCGGATGCCTACAGACCCCG
AGATGACAGCTTTTCATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGGAATAAAAC
TGGTGCGCAAGGTGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCGGTGTGGCA
TATCGACATGGTCAGCTTGAGGAGAATGACCGTGTTAGCCATCAATGGACATGATCTTCG
ATATGGCAGCCCAGAAAGTGCGGCTCATCTGATTGAGGCCAGTGAAAGACGTGTTACCTCG
TCGTGTCCCGCCAGGTTCCGGCAGCGGAGCCCTGACATCTTTCAGGAAGCCGGCTGGAACAGC
AATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACTCCCAAGCCCCCTCCATCCTAC
AATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCTCGGCATGA
CCGTGCGCAGGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCATCAGTGTTGAG
CCCGGAGGAGTCATAAGCAGAGATGGAAGAATAAAAACAGGTGACATTTTGTGTAATGTGGA
TGGGGTCGAACTGACAGAGGTGAGCCGGAGTGAGGCAGTGGCATTATTGAAAAGAACATCAT
CCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGCCCCAGGAAGACTGCAGCAGC
CCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGAAGTGGTCCCCATCCTGGGT
CATGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTACGAAGAAACA
CAGCTGGAAGTCTGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAAACAAACCT
TTTTTCATCAAATCCATTGTTGAAGGAACACCAGCATACAATGATGGAAGAATTAGATGTGG
TGATATTCTTCTGTGCTGTCAATGGTAGAAGTACATCAGGAATGATACATGCTTGCTTGGCAA
GACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTTTT
TTATAGAATCAATGATGGGTGAGAGGAAAACAGAAAAATCACAAATAGGCTAAGAAGTTGAA
ACACTATATTTATCTTGTCAGTTTTTTATATTTAAAGAAAGAATACATTGTAAAAATGTCAGG
AAAAGTATGATCATCTAATGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAATTA
AACTACTAGTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTACAACATTGTTTATATTT
TTTCTATTCAATAAAAAGCCCTAAAACACTAAAATGATTGATTGTATACCCCACTGAATT
CAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCA
TTTTTAATTTACAGCTAAATATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTCATCAA
ACAAGAATAAATATTTTTTCAGAAGTTAAA

150/310

FIGURE 147

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTAT
APSPEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRS
FKKINRALSVLRRTKSGSAVANHADQGRESENTTAPEVFPRLYHLIPDGEITSIKINRVDP
SESLSIRLVGGSETPLVHIIIIQHIIYRDGVIARDGRLLPGDIIILKVNGMDISNVPHNYAVRLL
RQPCQVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGV
FIFNVLDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQRQS
PDIFQEAGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHRE
WDLPIYVISVEPGGVISRDGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEV
KEYEPQEDCSSPAALDSNHNMAPPSDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIV
GGYEEYNGNKPFFIKSIVEGTPAYNDGRIRCGDILLAVNGRSTSGMIHACLARLLKELKGRI
TLTIVSWPGTFL

151/310

FIGURE 148

CCAAAGTGATCATTTGAAAAAGAGATATCCACATCTTCAAGCCCATATAAAGGATAGAAGCT
GCACAGGGCAGCTTTACTTACTCCAGCACCTTCCTCTCCCAGGCAAATGGTGCTGACCATCT
TTGGGATACAATCTCATGGATACGAGGTTTTTAACATCATCAGCCCAAGCAACAATGGTGGC
AATGTTCAAGGAGACAGTGACAATTGATAATGAAAAAATAACCGCCATCGTTAACATCCATGC
AGGATCATGCTCTTCTACCACAATTTTTGACTATAAACATGGCTACATTGCATCCAGGGTGC
TCTCCCGAAGAGCCTGCTTTATCCTGAAGATGGACCATCAGAACATCCCTCCTCTGAACAAT
CTCCAATGGTACATCTATGAGAAACAGGCTCTGGACAACATGTTCTCCAACAAATACACCTG
GGTCAAGTACAACCCTCTGGAGTCTCTGATCAAAGACGTGGATTGGTTCCTGCTGGGTCAC
CCATTGAGAACTCTGCAAACATATCCCTTTGTATAAGGGGGAAGTGGTTGAAAACACACAT
AATGTCGGTGCTGGAGGCTGTGCAAAGGCTGGGCTCCTGGGCATCTTGGGAATTTCAATCTG
TGCAGACATTCATGTTTTAGGATGATTAGCCCTCTTGTTTTATCTTTTCAAAGAAATACATCC
TTGGTTTACACTCAAAGTCAAATTAAATTCTTTCCCAATGCCCCAACTAATTTTGAGATTC
AGTCAGAAAATATAAATGCTGTATTTATA

FIGURE 149

MKILVAFLVVLTI FGIQSHGYEVFNIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT
IFDYKHGYIASRVLSRRACFILKMDHQNIPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLE
SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

153/310

FIGURE 150

GGCACGAGCCAGGAACTAGGAGGTTCTCACTGCCCCGAGCAGAGGCCCTACACCCACCGAGGC
ATGGGGCTCCCTGGGCTGTTCTGCTTGGCCGTGCTGGCTGCCAGCAGCTTCTCCAAGGCACG
GGAGGAAGAAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTTCCCCAAAG
GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC
TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT
CAACCTCAACGTCACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT
CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG
CCAGTGTCTGAGCTGCGGGCCAACTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGA
GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG
GGCAGGTCCACCTGCAGCAGAGACCATGCCACAGGCAGCCTGCCAACTTCTCCTTCCTGCCG
AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC
CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA
GCCCCATCCTTGCCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG
GGG TTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCCATGTAGAATGAACC
GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACTGTTTCGTATTTGGA
GTTTCATGCAAAATGAGTGTGTTTTAGCTGCTCTTGCCACAAAAAAAAAAAAAAAAAAAAA

FIGURE 151

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVPKGRWVLITCCAPQPPPPITYSL
CGTKNIKVAKKVVKTHEPASFNLNVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK
PVSELNANFTLQDRGAGPRVEMICQASSGSPITNSLIGKDGQVHLQQRPCHRQPANFSFLP
SQTSDWFWCQAANNANVQHSALTVVPPGGDQKMEDWQGPLESPILALPLYRSTRRLSEEEFG
GFRIGNGEVRGRKAAAM

155/310

FIGURE 152

GGTCCTTAATGGCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTG
CTGTCCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGT
CATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGA
CTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAAA
CTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACT
TACAGAGCAACTGCGTGACATTGAGCTGGAGAATTACACACCCCAAGGAACCCCTCACCCCTGC
AGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGT
TTCGATGGGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCC
TGGAGCCAGAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATT
ACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACC
CTGGAGCCAAGTGCAGGAGCACCCTCGCCATGTCCTCAGGCACAACCCAACTCAGGGCCAC
AGCCACCACCCTCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCA
TCTGAGGAGAGTCCTTTAGAGTGACAGGTAAAGCTGATACCAAAGGCTCCTGTGAGCAGC
GTCTTGATCAAACCTCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGT
GGCCTCCAGCAGATCATGATGACATCATGGACCCAATAGCTCATTCACTGCCTTGATTCCCTT
TTGCCAACAATTTTACCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACC
TGATGGAATTCCTGCACTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTC
TCTTTTTGTTTGGAAAATCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATT
GTCAGTAAATAATCACGTTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAG
AATTTTAAATTATTTAATAAGAAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTAT
TGTTCTGTACTGATATTTAAATAAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAA

156/310**FIGURE 153**

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTF
HYDCGNKTVTPVSPLGKKLVTTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQAR
MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS
MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

157/310

FIGURE 154

GGGAAAGCCATTTTCGAAAACCCATCTATACAAACTATATATTTTCATTTCTGCTGCTAGCTG
CCTTGGGCCTCACAATTTTCATTCTGTTTTCTGACTTTCAAGTTATATACCGTGGAATGGAG
TTGATCCCAACCATAACATCGTGGAGGGTTTTAATTTTGGTGGTAGCCCTCACCCAATTCTG
GTGTGGCTTTCTTTGCAGAGGATTCCACCTTCAAAATCATGAACTCTGGCTGTTGATCAAAA
GAGAATTTGGATTCTACTCTAAAAGTCAATATAGGACTTGGCAAAAGAAGCTAGCAGAAGAC
TCAACCTGGCCTCCCATAAACAGGACAGATTATTCAGGTGATGGCAAAAATGGATTCTACAT
CAACGGAGGCTATGAAAGCCATGAACAGATTCCAAAAAGAAACTCAAATTGGGAGGCCAAC
CCACAGAACAGCATTTCTGGGCCAGGCTGTAATCAGAATTGTGTCGTACATGCTCAACAGC
ATTGCTTTTTTCCCCAAAATTAACACATTGTGGAGAAGTGATGATACTCTCCCCTTACCTTT
CCTCTCTCCATTCAAGCATTCAAAGTATATTTTCAATGAATTAAACCTTGCAGCAAGGGACC
TTAGATAGGCTTATTCTGACTGTATGCTTTACCAATGAGAGAAAAAAATGCATTTCTGTAT
CATCCTTTTCAATAAACTGTATTCATTTTGAAAAAAAAAAAAAAAAAAAAA

158/310

FIGURE 155

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHFWLLIKREFGFYSKSQYRTWQKKLA
EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGQPTQHFWARL

FIGURE 156

GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACCATGAAG
CTCTTGTGGCAGGTAAGTGTGCACCACCACACCTGGAATGCCATCCTGCTCCCGTTCGTCTA
CCTCACGGCGCAAGTGTGGATTCTGTGTGCAGCCATCGCTGCTGCCGCCTCAGCCGGGCCCC
AGAACTGCCCCCTCCGTTTGTCTCGTGCAGTAACCAGTTTCAAGCAAGGTGGTGTGCACGCGCCGG
GGCCTCTCCGAGGTCCCGCAGGGTATTCCTCGAACACCCGGTACCTCAACCTCATGGAGAA
CAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCTGTCAGT
TGGGCAGGAACTCCATCCGGCAGATTGAGGTGGGGGCCCTTCAACGGCCTGGCCAGCCTCAAC
ACCCTGGAGCTGTTTCGACAAGTGGCTGACAGTCATCCCTAGCGGGGCCCTTGAATACCTGTC
CAAGCTGCGGGAGCTCTGGCTTCGCAACAACCCCATCGAAAGCATCCCCTCTTACGCCTTCA
ACCGGGTGCCCTCCCTCATGCGCCTGGACTTGGGGGAGCTCAAGAAGCTGGAGTATATCTCT
GAGGGAGCTTTTGAGGGGCTGTTCAACCTCAAGTATCTGAACTTGGGCATGTGCAACATTAA
AGACATGCCCAATCTCACCCCCCTGGTGGGGCTGGAGGAGCTGGAGATGTGAGGGAACCACT
TCCCTGAGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCTAG
AACTCACAGGTCAGCCTGATTGAGCGGAATGCTTTTGACGGGCTGGCTTCACTTGTGGAAGT
CAACTTGGCCCAATAACCTCTCTTCTTTGCCCATGACCTCTTTACCCCGCTGAGGTACC
TGGTGGAGTTGCATCTACACCACAACCCTTGGAACTGTGATTGTGACATTCTGTGGCTAGCC
TGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTGCTGCTCCCAT
GCACATGCGAGGCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCT
TCATCATGGACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAACTTAAGTGT
CGGACTCCCCCTATGTCTCCGTGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGC
CTCCCGCCACCCAAGGATCTCTGTCTCAACGACGGCACCTTGAAGTTCCTCCACGTGCTGC
TTTCAGACACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCAACTCCAACGCCTCG
GCCTACCTCAATGTGAGCAGGCTGAGCTTAACACCTCCAAGTACAGCTTCTTACCACAGT
AACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTCTTA
CCACGTCCACTGGTTACCAGCCGGCATATACCACCTCTACCACGGTGCTCATTACAGACTACC
CGTGTGCCCAAGCAGGTGGCAGTACCCGCGACAGACACCACTGACAAGATGCAGACCAGCCT
GGATGAAGTCATGAAGACCACCAAGATCATCATTGGCTGCTTTGTGGCAGTGACTCTGCTAG
CTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGCGGAGTACAGTC
ACAGCCGCCCCGACTGTTGAGATAATCCAGGTGGACGAAGACATCCCAGCAGCAACATCCGC
AGCAGCAACAGCAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCACCAATTC
ATGACCATATTAAGTACAAACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAACAGC
CTGGGGAAGTCTCTGCACCCCAAGTCAACCACTATCTCTGAACCTTATATAATTACAGCCCA
TACCAAGGACAAGGTACAGGAACTCAAATATGACTCCCTCCCCCAAAAACTTATAAAAT
GCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTTTTCTTGTA
TATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAACAGATTATATTAAATTTAAAGA
CAAAAAGTCAAAACA

FIGURE 157

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCT
RRGLSEVPQGI PSNTRYLNLMENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLAS
LNTLELFDNWLTVIPSGAFEYLSKLRELWLRNNPIESIPSYAFNRVPSLMRLDLGELKKLEY
ISEGAFEGLFNLKYLNLGMCNIKDMPNLTPLVGLEELEMMSGNHFPEIRPGSFHGLSSLKKLW
VMNSQVSLIERNAFDGLASLVELNLAHNNLSSLP HDLFTPLRYLVELHLHHPWNCD CDILW
LAWWLREYIPTNSTCCGRCHAPMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAEL
KCRTPPMSSVKWLLPNGTVLSHASRHPRI SVLNDGTLNFSHVLLSDTG VYTCMVTNVAGNSN
ASAYLNVSTAE LNTSNYSFFT TTVTVETTEISPEDTTRKYKVPVPTTSTGYQPAYTTSTTVLIQ
TTRVPKQVAVPATD TTDKMQTS LDEV MKTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQORS
TVTAARTVEIIQVDEDI PAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTE
NSLGNSLHPTVTTISEPYIIQHTTKDKVQETQI

162/310

FIGURE 158

[illegible]

162/310

FIGURE 159

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVV
GYTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGIFYCAECRAGWYGGD
CMRCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRD
GDNRDGQIIKRVCGNERPAPIQSIGSSLHVLHFHSDGSKNFDGFFHAIYEEITACSSSPCFHDG
TCVLDKAGSYKCACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTV
VSFFCNNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPQVQSRETPLH
QLYSAAFSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGK
WSGRAPSCIPICGKIENITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNE
RTVVAAHCVTDLGKVMTIKTADLKVVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLD
ADIAILKLLDKARISTRVQPICLAASRDLSTSFOESHITVAGWNVLADVRSFGKNDTLRSG
VVSVDSSLCEEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPGRASPEPR
WHLMGLVSWSYDKTCSHRLSTAFTKVLPFKDWIERNMK

163/310

FIGURE 160

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGA
AGCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGC
TTCAACCTGACTTTCCACCTTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCT
GACAGTGGTTGGGTGGGCCACCAGTAACTACTTTCGTGGGTGCCATTCAAGAGATTCCTAAAG
CAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACT
AATGAAGCATCCACGAAGAAGGTAGAACTTGACAACGTCTCTTCTGTGTCTCCTTACCTCAG
AGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATC
CCAAAGTGTCCAGAGGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATCCTC
GTTCCCCACCGGAACAGAGAGAAAACCTGATGTACCTGCTGGAACATCTGCATCCCTTCCT
GCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGTTTA
ATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGACTGC
TTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGA
GCATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGAT
ATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTTCAAGGTGAATGGATTCTCTAAC
AACTACTGGGGATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAAT
GAAAATTTCCCGGCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACA
AAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGA
ACAGATGGGTTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATATAT
CAACATCACAGTGGATTTCTGGTTTGGTGCATGACCCTGGATCTTTTGGTGATGTTTGGGAAG
AACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTAAGA
ACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCTTTTGTATTTTCTTAGCAGAGCTCCT
GGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGATCATGAGG
GTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAAA
ATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTAT
GGGATAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGT
CCAAGGTAGAAAGGTACGAAGATACAATACTGTTATTTCATTTATCCTGTACAATCATCTGTG
AAGTGGTGGTGTGAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGA
CACAGTGAACCTGGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTGCGCTGCAAAGGCAGCAG
TAGCTGAGCTGGTTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCTTCC
AGTGATGCCACCAGAGAATACATTCTCTATTAGTTTTTTAAAGAGTTTTTGTAAAATGATTT
TGTAACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACATATTAATAATAAATATGT
CTATCAAATACCTCTGTAGTAAATGTGAAAAAGCAAAA

164/310

FIGURE 161

MGFNLT FHLSYKFRL L L L L L T L C L T V V G W A T S N Y F V G A I Q E I P K A K E F M A N F H K T L I L G K G K T
L T N E A S T K K V E L D N C P S V S P Y L R G Q S K L I F K P D L T L E E V Q A E N P K V S R G R Y R P Q E C K A L Q R V
A I L V P H R N R E K H L M Y L L E H L H P F L Q R Q Q L D Y G I Y V I H Q A E G K K F N R A K L L N V G Y L E A L K E E N
W D C F I F H D V D L V P E N D F N L Y K C E E H P K H L V V G R N S T G Y R L R Y S G Y F G G V T A L S R E Q F F K V N G
F S N N Y W G W G G E D D L R L R V E L Q R M K I S R P L P E V G K Y T M V F H T R D K G N E V N A E R M K L L H Q V S R
V W R T D G L S S C S Y K L V S V E H N P L Y I N I T V D F W F G A

Important features:**Signal peptide:**

amino acids 1-27

N-glycosylation sites:

amino acids 4-7, 220-223 and 335-338

Xylose isomerase proteins:

amino acids 191-201

165/310

FIGURE 162

CGTGGGCCGGGGTCCGCGCAGCGGGCTGTGGGCGCGCCCCGGAGGAGCGACCGCCGCAGTTCTC
GAGCTCCAGCTGCATTCCCTCCGCGTCCGCCCCACGCTTCTCCCGCTCCGGGCCCCGCAATG
GCCCAGGCAGTGTGGTCCGCGCCTCGGCCGCATCCTCTGGCTTGCCTGCCTCCTGCCCTGGGC
CCCGGCAGGGGTGGCCGCAGGCCTGTATGAACTCAATCTCACCACCGATAGCCCTGCCACCA
CGGGAGCGGTGGTGACCATCTCGGCCAGCCTGGTGGCCAAGGACAACGGCAGCCTGGCCCTG
CCCGCTGACGCCACCTCTACCGCTTCCACTGGATCCACACCCCGCTGGTGTCTACTGGCAA
GATGGAGAAGGTCTCAGCTCCACCATCCGTGTGGTGGGCCACGTGCCCGGGGAATTCCCGG
TCTCTGTCTGGGTCACTGCCGCTGACTGCTGGATGTGCCAGCCTGTGGCCAGGGGCTTTGTG
GTCTCCCCATCACAGAGTTCCTCGTGGGGGACCTTGTTGTCACCAGAACAACCTTCCCTACC
CTGGCCCAGCTCCTATCTCACTAAGACCGTCTGAAAGTCTCCTTCCCTCCACGACCCGA
GCAACTTCTCAAGACCGCTTGTTTCTCTACAGCTGGGACTTCGGGGACGGGACCCAGATG
GTGACTGAAGACTCCGTGGTCTATTATAACTATTCCATCATCGGGACCTTCACCGTGAAGCT
CAAAGTGGTGGCGGAGTGGGAAGAGGTGGAGCCGGATGCCACGAGGGCTGTGAAGCAGAAGA
CCGGGGACTTCTCCGCCTCGCTGAAGCTGCAGGAAACCTTCGAGGCATCCAAGTGTGGGG
CCCACCCTAATTAGACCTTCCAAAAGATGACCGTGACCTTGAACCTTCTGGGGAGCCCTCC
TCTGACTGTGTGCTGGCGTCTCAAGCCTGAGTGCCTCCCGCTGGAGGAAGGGGAGTGCCACC
CTGTGTCCGTGGCCAGCACAGCGTACAACCTGACCCACACCTTCAGGGACCTGGGGACTAC
TGCTTCAGCATCCGGGCCGAGAATATCATCAGCAAGACACATCAGTACCACAAGATCCAGGT
GTGGCCCTCCAGAATCCAGCCGGCTGTCTTTGCTTTCCCATGTGCTACACTTATCACTGTGA
TGTTGGCCTTCATCATGTACATGACCTGCGGAATGCCACTCAGCAAAAGGACATGGTGGAG
AACCCGGAGCCACCCTCTGGGGTCAGGTGCTGCTGCCAGATGTGCTGTGGGCCTTTCTTGCT
GGAGACTCCATCTGAGTACCTGGAAATTGTTCTGTGAGAACCACGGGCTGCTCCCGCCCCTCT
ATAAGTCTGTCAAACTTACACCGTGTGAGCACTCCCCCTCCCCACCCCATCTCAGTGTTAA
CTGACTGCTGACTTGGAGTTCCAGCAGGGTGGTGTGCACCACTGACCAGGAGGGGTTCAAT
TGCGTGGGGCTGTTGGCCTGGATCATCCATCCATCTGTACAGTTCAAGCACTGCCACAAGCC
CCTCCCTCTCTGTACCCCTGACCCAGCCATTACCCATCTGTACAGTCCAGCCACTGACA
TAAGCCCCACTCGGTTACCACCCCTTGACCCCTACCTTTGAAGAGGCTTCGTGCAGGACT
TTGATGCTTGGGGTGTTCGTGTTGACTCCTAGGTGGGCCTGGCTGCCCACTGCCCATTCCT
CTCATATTGGCACATCTGCTGTCCATTGGGGGTTCTCAGTTTCTCCCCAGACAGCCCTAC
CTGTGCCAGAGAGCTAGAAAGAAGGTCATAAAGGGTTAAAAATCCATAACTAAAGGTTGTAC
ACATAGATGGGCACACTCACAGAGAGAAGTGTGCATGTACACACACCACACACACACACA
CACACACACACAGAAATATAAACACATGCGTCACATGGGCATTTAGATGATCAGCTCTGTA
TCTGGTTAAGTCGGTTGCTGGGATGCACCCTGCACTAGAGCTGAAAGGAAATTTGACCTCCA
AGCAGCCCTGACAGGTTCTGGGCCCCGGGCCCTCCCTTTGTGCTTTGTCTCTGCAGTTCTTGC
GCCCTTTATAAGGCCATCCTAGTCCCTGCTGGCTGGCAGGGGCTGGATGGGGGGCAGGACT
AATACTGAGTGATTGCAGAGTGCTTTATAAATATCACCTTATTTTATCGAAACCCATCTGTG
AACTTTTCACTGAGGAAAAGGCCTTGCCAGCGGTAGAAGAGGTTGAGTCAAGGCCGGGCGCGG
TGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACGAGATCAGGA
GATCGAGACCACCCTGGCTAACACGGTGAACCCCGTCTCTACTAAAAAATACAAAAAGTT
AGCCGGGCGTGGTGGTGGGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATG
GTGCGAACCCGGGAGGCGGAGCTTGCAGTGAGCCAGATGGCGCCACTGCACTCCAGCCTGA
GTGACAGAGCGAGACTCTGTCTCCA

166/310

FIGURE 163

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA
LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF
VVLPITEFLVGDLVVTQNTSLPWPSYLTKTVLKVSFLLHDPSNFLKTALFLYSWDFGDGTQ
MVTEDSVVYYNYSIIGTFTVKLVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL
GPTLIQTFQKMTVTNLNFGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLTHTFRDPGD
YCFSIRAENIISKTHQYHKIQVWPSRIQPAVFAPPCATLITVMLAFIMYMTLRNATQOKDMV
ENPEPPSGVRCCCQMCCGPFLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domain:

amino acids 339-362

N-glycosylation sites.

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

168/310

FIGURE 165

MALSSQIWAACLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTH
FPICIFCCGCCHRSKCGMCCKT

169/310

FIGURE 166

CTGTCAGGAAGGACCATCTGAAGGCTGCAATTTGTTCTTAGGGAGGCAGGTGCTGGCCTGGC
CTGGATCTTCCACCATGTTCTGTTGCTGCCTTTTGATAGCCTGATTGTCAACCTTCTGGGC
ATCTCCCTGACTGTCCTCTTCACCCTCCTTCTCGTTTTTCATCATAGTGCCAGCCATTTTTGG
AGTCTCCTTTGGTATCCGCAAACCTCTACATGAAAAGTCTGTTAAAAATCTTTGCGTGCGGCTA
CCTTGAGAATGGAGCGAGGAGCCAAGGAGAAGAACCACCAGCTTTACAAGCCCTACACCAAC
GGAATCATTGCAAAGGATCCCACCTTCACTAGAAGAAGAGATCAAAGAGATTTCGTGCAAGTGG
TAGTAGTAAGGCTCTGGACAACACTCCAGAGTTCGAGCTCTCTGACATTTTCTACTTTTGGC
GGAAGGAATGGAGACCATTTATGGATGATGAGGTGACAAAGAGATTCTCAGCAGAAGAACTG
GAGTCCTGGAACCTGCTGAGCAGAACCAATTATACTTCCAGTACATCAGCCTTCGGCTCAC
GGTCCTGTGGGGTTAGGAGTGCTGATTCCGTACTGCTTTCTGCTGCCGCTCAGGATAGCAC
TGGCTTTCACAGGGATTAGCCTTCTGGTGGTGGGCACAACTGTGGTGGGATACTTGCCAAAT
GGGAGGTTTAAGGAATTCATGAGTAAACATGTTCACTTAATGTGTTACCGGATCTGCGTGCG
AGCGCTGACAGCCATCATCACCTACCATGACAGGGAAAACAGACCAAGAAATGGTGGCATCT
GTGTGGCCAATCATACCTCACCGATCGATGTGATCATCTTGGCCAGCGATGGCTATTATGCC
ATGGTGGGTCAAGTGACGGGGGACTCATGGGTGTGATTGAGAGAGCCATGGTGAAGGCCTG
CCCACACGTCTGGTTTGAGCGCTCGGAAGTGAAGGATCGCCACCTGGTGGCTAAGAGACTGA
CTGAACATGTGCAAGATAAAAGCAAGCTGCCTATCCTCATCTTCCAGAAGGAACCTGCATC
AATAATACATCGGTGATGATGTTCAAAAAGGGAAAGTTTTGAAATTGGAGCCACAGTTTACCC
TGTTGCTATCAAGTATGACCCTCAATTTGGCGATGCCTTCTGGAACAGCAGCAAATACGGGA
TGGTGACGTACCTGCTGCGAATGATGACCAGCTGGGCCATTGTCTGCAGCGTGTGGTACCTG
CCTCCCATGACTAGAGAGGCAGATGAAGATGCTGTCCAGTTTGCGAATAGGGTGAAATCTGC
CATTGCCAGGCAGGGAGGACTTGTGGACCTGCTGTGGGATGGGGGCCTGAAGAGGGAGAAGG
TGAAGGACACGTTCAAGGAGGAGCAGCAGAAGCTGTACAGCAAGATGATCGTGGGGAACCA
AAGGACAGGAGCCGCTCCTGAGCCTGCCTCCAGCTGGCTGGGGCCACCGTGCGGGGTGCCAA
CGGGCTCAGAGCTGGAGTTGCCGCCGCCGCCCCACTGCTGTGTCTTTCCAGACTCCAGGG
CTCCCCGGGCTGCTCTGGATCCCAGGACTCCGGCTTTCGCCGAGCCGCAGCGGGATCCCTGT
GCACCCGGCGCAGCCTACCCTTGGTGGTCTAAACGGATGCTGCTGGGTGTTGCGACCCAGGA
CGAGATGCCTTGTTTCTTTTACAATAAGTCGTTGGAGGAATGCCATTAAAGTGAACCTCCCA
CCTTTGCACGCTGTGCGGGCTGAGTGGTTGGGGAGATGTGGCCATGGTCTTGTGCTAGAGAT
GGCGGTACAAGAGTCTGTTATGCAAGCCCGTGTGCCAGGGATGTGCTGGGGGCGGCCACCCG
CTCTCCAGGAAAGGCACAGCTGAGGCACTGTGGCTGGCTTCGGCCTCAACATCGCCCCCAGC
CTTGAGACTCTGCAGACATGATAGGAAGGAACTGTCTATCTGCAGGGGCTTTAGCAAATG
AAGGGTTAGATTTTTATGCTGCTGCTGATGGGGTTACTAAAGGGAGGGGAAGAGGCCAGGTG
GGCCGCTGACTGGGCCATGGGGAGAACGTGTGTTTCGTACTCCAGGCTAACCTGAACTCCCC
ATGTGATGCGCGCTTTGTTGAATGTGTGCTCGGTTTCCCATCTGTAATATGAGTCGGGGG
GAATGGTGGTGATTCCCTACCTCACAGGGCTGTTGTGGGGATTAAAGTGCTGCGGGTGAGTGA
AGGACACATCACGTTTCAAGTACAGGCCCAAAAACGGGGCACGGCAGGCCTGAG
CTCAGAGCTGCTGCACTGGGCTTTGGATTTGTTCTTGTGAGTAAATAAACTGGCTGGTGAATGA

170/310

FIGURE 167

MFLLLPFDSLIVNLLGISLTVLFTLLLVIIVPAIFGVVSFGIRKLYMKSLLKIFAWATLRME
RGAKEKNHQLYKPYTNGIIAKDPTSLEEEIKEIRRS GSSKALDNTPEFELSDIFYFCRKGME
TIMDDEVTKRFSABEELESWNLLSRTNYNFQYISLRLTVLWGLGVLI RYCFL LPLRIALFTG
ISLLVVGTTVVGYPNGRFKEFMSKHVHLMCYRICVRALTAIITYHDRENRP RGGICVANH
TSPIDVIIILASDGY YAMVGQVHGGLMGVIQ RAMVKACPHVWFERSEVKDRHLVAKRLTEHVQ
DKSKLPILIFPEGTCINNTSVMMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMVTYL
LRMMTSWAIVCSVWYLPMTREADEDAVQFANRVKSAIARQGGLVDLLWDGGLKREKVKDTF
KEEQQKLYSKMIVGNHKDRSRS

171/310
FIGURE 168

GCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCA
CGTCTCCTCCAGGGATGGCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCAC
ACCTGGCAGGCCAGGCTGTTCCCACCATCCTGCCCCCTGGGCCTGGCTCCAGACACCTTTGA
CGATACCTATGTGGGTTGTGCAGAGGAGATGGAGGAGAAGGCAGCCCCCTGCTAAAGGAGG
AAATGGCCCCACCATGCCCTGCTGCGGGAATCCTGGGAGGCAGCCAGGAGACCTGGGAGGAC
AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT
CTACACCAACTCATCGAACACCTTGTAAGTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG
GCTCCCGGGAGCTCTACATGAGGCACTTTCCTTCAAGGCCCTGCATTTCTACCTGATCCGG
GCCCTGCAGCTGCTGCGAGGCAGTGGGGGCTGCAGCAGGGGACCTGGGGAGGTGGTGTTCGG
AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT
TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCACAGATTTGGGGAGAAGAGGCGGGGCTGT
GTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCTG
GAAGACTCTGCTCTTGCCCCCTGGAGAGTTCCAGCTCTCAGGGGTGGGGCCCTGAAAGTCCA
ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG
CAGCCTTGAGAAGCAAGAACATGGTTCGGGACCCAGCCCTAGCAGCCTTCTCCCCAACCAGG
ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACCTCTGCTATGTGATGGGGACTTCCT
GGGACAAGCAAGGAAAGTACTGAGGCAGCCACTTGATTGAACGGTGTTGCAATGTGGAGACA
TGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA

172/310

FIGURE 169

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEEKAAPLLKEEMAHH
ALLRESWEAAQETWEDKRRGLTLPFGFKAQNGIAIMVYTNSNTLYWELNQAVRTGGGSREL
YMRHFPPKALHFYLIRALQLLRGSGGCSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASSS
LDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPPWKTLLLAPGEFQLSGVGP

173/310
FIGURE 170

GTGGCTTCATTTAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCA
CCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTG
GTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTC
TATTGTCTGGACCTTCAACACAACCCCTCTTGTACCATACAGCCAGAAGGGGGCACTATCA
TAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAG
CTCAGCAAATGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACT
CCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAG
TCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATG
GAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTC
CCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCT
GCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGT
GAAGGTGCTGCTGATGACCCAGATTCTCCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCCCT
CCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTCTGAAGAGAGAGAGACAAGAAG
AGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCTGGGAAACTCCTAACATATGCCCCCAT
TCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGA
TCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGATGGAAATCCCCACTCAC
TGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTATCTAGACAGCAGTG
CACTCCCCTAAGTCTCTGCTCA

174/310

FIGURE 171

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVT
IQPEGGTIIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHV
YEHLSPKPKVTMGLQSNKNGTCVTNLTCCMEHGEEDVIYTWKALGQAANESHNGSILPISWRW
GESDMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLW
FLKRERQEEYIEKKRVDICRETPNICPHSGENTHEYDTIPHTNRTILKEDPANTVYSTVEIP
KKMENPHSLLTMPDTPRLFAYENVI

175/310
FIGURE 172

CTGGTTCCCCAACATGCCTCACCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCC
TCTGGACCCGTGAAAGAGCTGGTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTC
CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCCACATAC
AGCCAGAAGGGGGCACTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA
GATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGT
GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG
AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG
ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT
GGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG
AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCC
ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATGGTCCTCCT
GTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTAAGTGGGGCTATTTCTTTGGTTTC
TGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCTGGGAA
ACTCCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAA
TAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA
AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAG
AATGTTATCTAGACAGCAGTGCACTCCCCTAAGTCTCTGCTCAAAAAAAAAAAAAAAAAAAAA

176/310

FIGURE 173

GAAAGACGTGGTCCTGACAGACAGACAATCCTATTCCTACCAAAATGAAGATGCTGCTGCT
GCTGTGTTTGGGACTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGAAGGA
ACTTTAATGTAGAAAAGATTAATGGGGAATGGCATACTATTATCCTGGCCTCTGACAAAAGA
GAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTGGAGCAAATCCATGTCTTGGAGAA
TTCCTTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTATCTATGGTTG
CTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATTCAATACATTTACT
ATACCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGA
AACCTTCCAGCTGATGGGGCTCTATGGCCGAGAACCAGATTTGAGTTCAGACATCAAGGAAA
GGTTTGCACAACATATGTGAGGAGCATGGAATCCTTAGAGAAAATATCATTGACCTATCCAAT
GCCAATCGCTGCCTCCAGGCCCGAGAATGAAGAATGGCCTGAGCCTCCAGTGTTGAGTGGAC
ACTTCTCACCAGGACTCCACCATCATCCCTTCCTATCCATACAGCATCCCCAGTATAAATTC
TGTGATCTGCATTCCATCCTGTCTCACTGAGAAGTCCAATTCCAGTCTATCAACATGTTACC
TAGGATACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTCTTTGATACACCCTTGACAAT
TTTTCATGAAATTATTCCTCTTCCTGTTCAATAAATGATTACCCTTGCACTTAA

177/310

FIGURE 174

MKMLLLLCLGLTLVCVHAEASSTGRNFNVEKINGEWHTI ILASDKREKIEEHGNFRLFLEQ
IHVLENSLVCLKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLI
NEKDGETFQLMGLYGREPDLSSDIKERFAQLCEEHGILRENI IDLSNANRCLQARE

178/310

FIGURE 175

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAA
TGGATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTG
TCAGCTTAGTTGAGGAAGACCAATTTTCTCAAAACCCCATCTCTTGCTTTGAGTGGTGGTTC
CCAGGAATTATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAG
AAAAAGAGCGTGCTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCA
CAGTCATTGGTGCTCTGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTC
ATGTGTAATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGA
CATTCAATCCAGAATCCTTCAACTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTG
GTTTCAATAAACCCACCAGTAACGACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCAC
TTCGATTCTGAAGAAAACAAACATAGGCTTATCCACTTCTCAGTATTTTTAGGTCTATTGCT
TGTTGGAATTCTGGAGGTCCTGTTTGGGCTCAGTCAGATAGTCATCGGTTTTCCTTGGCTGTC
TGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTGTGTAGTTTAATGGGAATAAAATGTAAGTA
TCAGTAGTTTGAAAAAAAAAAAA

179/310

FIGURE 176

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA
IPATTMSLTARKKRACNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNA
NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRL
IHFSVFLGLLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

180/310

FIGURE 177

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACCATGAGGCT
GTCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCT
GCCCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTC
CAAGTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTG
CACCGATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAATAGTGAA
AAAATGTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCAAAGTCTTTCAACGACACC
CTGATCTTCACTAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

184/310

FIGURE 178

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNQLQVAKLNPPPEALAAKLEV
KHCTDQISFKKRLSLKKSWWK

182/310
FIGURE 179

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGG
TGAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTG
GAGCAGATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTC
GAACTGTGACATGGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGG
AAGCCAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAG
CTGAGCGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGG
CAAATGCAAATACAAGAGCAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCAC
TCATCACTCCAGGCTCTGCCACTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGA
AGCCTAACACTGGCCCCCAGCACCTCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTC
TCTCCAAGGGCAGGCTGTTAGGCCCTTTCTGATCAGGAGGCTTCTTTATGAATTAAACTCG
CCCCACCACCCCCTCA

—183/310

FIGURE 180

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCK
YKSSQKQHSPVPEKAIPITPGSATTC

184/310

FIGURE 181

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCC
TGGCTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTC
CTGGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCC
AGTGTTTTCCACAGCCCCCAAACGGAACTGGTTTTGGGGTCACCTGGGCCTGATCACTCCT
ACAGAGGAGGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGT
ATGGCTGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCA
CCAATGCCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGCCCTGG
CTGGGAGAAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGAC
GCCCCGCTTCCATTTCAACATCCTGAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACA
TCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAG
CACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTG
TCAGGAGAGGCCCAGTGAATATATTGCCACCATCTTGGAGCTCAGTGCCCTTGTAAGAAAA
GAAGCCAGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGGCGGCGC
TTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGACGCTGTCATCCGGGAGCGGCGTCG
CACCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGG
ATTTCAATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTGATGAGGAT
ATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTC
CTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGC
AAGAGCTTCTGAAGGACCGCATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCC
TTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATCTCCCG
ATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTGCC
TCATCGATATTATAGGGGTCCATCACAACCCAACCTGTGTGGCCGGATCCTGAGGTCTACGAC
CCCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCTTTCTC
CGCAGGGCCCAGGAAGTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCCTGG
CGTTGATGCTGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAA
TTGATCATGCGCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCA
GTGACTTTCTGACCCATCCACCTGTTTTTTTGCAGATTGTGATGAATAAAACGGTGCTGTCAA

-185/310

FIGURE 182

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWG
HLGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF
IRFLKPWLGE GILLSGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQH LASEGS
SRLDMFEHISLMTLDSLQKCIFSFD SHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY
LSHDGRRFHRACRLVHDFTDAVIRERRRTLPTQGIDDFKDKAKSKTLD FIDVLLLSKDEDG
KALSDEDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEW
DDLAQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

186/310

FIGURE 183

CAACAGAAGCCAAGAAGGAAGCCGTCTATCTTGTGGCGATCATGTATAAGCTGGCCTCCTGC
TGTTTGCTTTTCACAGGATTCTTAAATCCTCTCTTATCTCTTCCTCTCCTTGACTCCAGGGA
AATATCCTTTCAACTCTCAGCACCTCATGAAGACGCGCGCTTAACTCCGGAGGAGCTAGAAA
GAGCTTCCCTTCTACAGATATTGCCAGAGATGCTGGGTGCAGAAAGAGGGGATATTCTCAGG
AAAGCAGACTCAAGTACCAACATTTTTTAACCCAAGAGGAAATTTGAGAAAGTTTCAGGATTT
CTCTGGACAAGATCCTAACATTTTACTGAGTCATCTTTTGGCCAGAATCTGGAAACCATACA
AGAAACGTGAGACTCCTGATTGCTTCTGGAAATACTGTGTCTTGAAGTGAAATAAGCATCTGT
TAGTCAGCTCAGAAACACCCATCTTAGAATATGAAAAATAACACAATGCTTGATTTGAAAAC
AGTGTGGAGAAAAACTAGGCAAACCTACACCCTGTTTCATTGTTACCTGGAAAATAAATCCTCT
ATGTTTTGCACAAAAAAAAAAAAAAAAA

187/210

FIGURE 184

MYKLASCCLLFTGFLNPLLSLPLLDREISFQLSAPHEDARLTPEELERASLLQILPEMLGA
ERGDILRKADSSTNIFNPRGNLRKFQDFSGQDPNILLSHLLARIWKPYKKRETPDCFWKYCV

-188/310

FIGURE 185

GAACATTTT TAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGAT
GGGGTTGCTGGTTTAAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCA
CCACCTCCGCCAGGAAGTGCAGGCCCCACCTGTCTGCAACCCAGCTGAGGCCATGCCCTCCCC
AGGGACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCT
CCAGCTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTCGAAGAAGCCACCA
GCCAAGCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCCGGAAGATGGAGGTCAAGCAGA
AGGGGCAGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGT
CAGGGGTT CAGTACCAGCAGCACAGCCAGGCCCTGGGGAAGTTTCTTCAGGACATCCTCTGG
GAAGAGGCCAAAGAGGCCCCAGCCGACAAGTGATCGCCCCACAAGCCTTACTCACCTCTCTCT
AAGTTTAGAAGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTA
CAAGCTCAGGAGGCGAATAAATGTTCAAAGTGTA

189/310

FIGURE 186

MPSPGTVCSLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG
GQAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

190/310

FIGURE 187

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTCCTCGTCCAGTACCTC
GTGAACCCCGGGGTGCTCCGCACGGACCCCAAGATGTCAAGAATATGAACACGTGGCTGCTGT
TCCTCCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTG
CTCCTGCTGGACTTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCT
GAGTATGTCCCCACCCTAAGCCCCCGATCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATC
TTACACCTCTACTTGAGTATGTCCCTAACCCCTGAGCCCCCACGCCTGGGGCCAGAGTCTTT
GTCCCCCGTGTGCGCATGTGTTTCAGGGTCAGCCTCTCCAGAAAGTGAGATCATGGACAAAAA
GGGCAAATCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCCCAGCAAGAAGCTGAAC
TCACGCCGAGACCTGCAGGAGTGGTGCCAGGTGCTTGAAGTAACAAGTTTAAATGTTTCTGAG
GACAATGGAATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTT
CCAAAAACACAAGTAGAAATTCTAACAATGAAATATATTACAGGCAGGTACCCACTAACCA
AACAACCTGAAGCGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTC
AGTCATGTTGCTGAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGT
AACAACAACCTCCCTGCTCCTGGCACCAGCCGTTTTTGGTCATGGTGGGCCAGCTGCAAAGCG
TCTTCCATTCTCTGGGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTCTGTGGAC
ACGGGCAGCAGAGTGTGTCCAGGCCAGCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCA
ACCCCTGGTCAGGGCAGAGGGAGTTGGGTGGGTGAGGCTCTGGGCTCACCTCCATCTCCAGA
GCATCCCCCTGCCTGCAGTTGTGGCAAGAACGCCAGCTCAGAATGAACACACCCCCACCAAGA
GCCTCCTTGTTTATAACACAGGTTACCCTACAAACCACTGTCCCCACACAACCTGGGGAT
GTTTTAAACACACACCTCTAACGCATATCTTACAGTCACTGTTGTCTTGCCTGAGGGTTGA
ATTTTTTTTAAATGAAAGTGCAATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

194/310

FIGURE 188

MNTWLLFLPLFPVQVQTLIVVVIIGMLVLLLDLGLVHLGQLLIFHIYLSMSPTLSPRSPQGW
VVRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQ
AQQEAELTPRPAGVVPGA

192/310

FIGURE 189

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACCATGGCCAAG
ATGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT
ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGG
TGCCCAAGCCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTG
GATGGAGATAACCAACACATCCACCCAGGAGGTGGTACAATACTGGGAGACTGGGGATGA
CCGGTTCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAACTGTGGAAG
AACCAGGGGAGAGGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAA
GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA
GCGGTTGATGGAGAAGGCTTCCCTCCCCTCCCCTCCCTTGGGGCTTTGTGGCAAAAATCCTA
TGGTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGCTTCCTCCTGCT
ACTAACAGACTTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAACTGAGCGCCTTTGCTG
CTGTTTCCTCTGTCCTGTCAGGTCTCCTGGGGATGGTGGCCACATGATGTATTCAACAAGTC
TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAATTATGGCTG
GGCCTTCTACATGGCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCAACCACCTTCA
ACACGTACACCAGGATGGTGCTGGAGTTCAAGTGCAAGCATAGTAAGAGCTTCAAGGAAAAC
CCGAAGTGCCTACCACATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCAC
CGTGGGTCCCTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG
GAGTCGACTTCTACTCCGAGCTGCGGAACAAGGGATTTCAAAGAGGGGCCAGCCAGGAGCTG
AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA
GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG
TCTCTTGAGCATGGTTTTTTAGAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC
CTAAGGGATTCTTGGGTGCCACTGCTCTCTTTTCTCTACAGCTCCATCTTGTTTCACCCAC
CCCACATCTCACACATCCAGAATTCCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGC
TAAACCATGGAGATAAAAAGAAGAGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

193/310
FIGURE 190

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFGVTQKVPKPLCEKGLAAKCFDMP
VSLDGDNTNTSTQEVVQYNWETGDDRF SFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKR
GEKGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQL
PPATNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSPDCQLGSRRLETTCLE
LWLGLLHGLALLHLLHGVGCHHLQHVVHODGAGVQVQA

124/310
FIGURE 191

AACTGGAAGGAAAGAAAGAAAGGTCAGCTTTGGCCCAGATGTGGTTACCCCTTGGTCTCCTG
TCTTTATGTCTTTCTCCTCTTCCTATTCTGTTCATCTCCCTCACTTAAGTCTCAGGCCTGTCA
GCAGCTCCTGTGGACATTGCCATCCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATG
TTATGGATGTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGT
GATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAAATCTGGCATG
AGATGGCACAGGTGACCACGCAGAAGCCACCAGAATCTTGCCTGCCCTATTCCTCCTCCCAA
GTCTGTTCTCTTATTGTCAACCTCAGCACAACAGGCTGGCGCCAATGGCATTACAGAGAAAG
CAATCTGTGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGT
AGCCACCTCCCTGTGAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCCGCCGTAGATTGAG
GACATTGCCCCCTGTGTGCCACCAACCAGGACTTTCCTTGGCTTGGCATCCCTGGCTCT
CTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACT
ATGGCGATGGCCATGATGTTACAATCCCACTTGCCTGAATAATCAAGTGGGAAGGGGAAGCA
GAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAAA
GGGAAGCAACAGGAACCTTCTGCAACTGGTTTTTATCGGAAAGATCATCCTGCCTGCAGATGC
TGTTGAAGGGGCACAAGAAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAA
ATAGAACAGTCTGCTGGGAGTCAGACCTGGAATTCTGATTCCAAACTCTTTATTACTTTGGG
AAGTCACTCAGCCTCCCCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGG
AACCAAGGAAACTAACAATGTAGGTTACTAGTGAATACCCCAATGGTTTCTCCAATTATGCC
CATGCCACCAAAACAATAAAACAAAATTCTCTAACACTGAAA

135/310

FIGURE 192

MWLPLGLLSLCLSPLPILSSPSLKSQACQQLLWTLPSPLVAFRANRTTYVMDVSTNQSGME
HRNHLCFCDLYDRATSPPLKCSLL

-195/310

FIGURE 193

GTAGCGCGTCTTGGGTCTCCCGGCTGCCGCTGCTGCCGCCGCCGCTCGGGTCTGGAGCCA
GGAGCGACGTCACCGCCATGGCAGGCATCAAAGCTTTGATTAGTTTGTCTTTGGAGGAGCA
ATCGGACTGATGTTTTTGTATGCTTGGATGTGCCCTTCCAATATACAACAAATACTGGCCCT
CTTTGTTCTATTTTTTTACATCCTTTACCTATTCCATACTGCATAGCAAGAAGATTAGTGG
ATGATACAGATGCTATGAGTAACGCTTGTAAGGAACCTGCCATCTTTCTTACAACGGGCATT
GTCGTGTCAGCTTTTGGACTCCCTATTGTATTGTCAGAGCACATCTGATTGAGTGGGGAGC
TTGTGCACTTGTCTCACAGGAAACACAGTCATCTTTGCAACTATACTAGGCTTTTCTTGG
TCTTTGGAAGCAATGACGACTTCAGCTGGCAGCAGTGGTGAAGAAATTAAGTAATATTG
TCAAATGGACTTCCTGTCAATTGTTGGCCATTACGCACACAGGAGATGGGGCAGTTAATGC
TGAATGGTATAGCAAGCCTCTTGGGGGTATTTTAGGTGCTCCCTTCTCACTTTTATTGTAAG
CATACTATTTTCACAGAGACTTGCTGAAGGATTAAAAGGATTTTCTCTTTTGGAAAAGCTTG
ACTGATTTCACTTATCTATAGTATGCTTTTGTGGTGTCTGCTGAATTTAAATATTTAT
GTGTTTTCTCTGTTAGGTGATTTTTTTTGGAAATCAATATGCAATGTTAAACACTTTTTTAA
TGTAATCATTTGCATTGGTTAGGAATTCAGAATTCGCCCGGCTCTATTACTGGTCAAGTACA
TCTTTTCTCTTAAATATTTAGCCTCCATTATTACAAAAAATTATAAAAAATAAGTTTTTCA
TCAGTCAGGATGACATCACTCCCAATGTTATGCAGACATACAGACGGTGGCATAACGTTATA
GACTGTATACTCAGTGCAAATATAGCTGCATTTATACCTCAGAGGGGCCAAGTGTTAATGCC
CATGCCCTCCGTTAAGGGTTGTTGGTTTTACTGGTAGACAGATGTTTTGTGGATTGAAAATT
ATTTTATGGAATTGCTACAGAGGAGTGCTTTTCTCTCAATTGTTAGAAGAATTTATGTTAA
ACTTTAAGGTAAGGGTGTAACAAACATTTTTGAGATAAGGTTTTATTTATGTTTATTATTGT
TAGAGTGAGTTGCAATGTGGGAAGAAATGACATTGAAATTCAGTTTTTGAATCCTGTTTCT
ATTTATAAGTGAAATTTGTGATCTCCTATCAACCTTTTCATGTTTTACCCTGTTAAATGGAC
ATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTTGCATCATATATGCCAGAAAACC
TTCCTCTGCTTCCTCCTTTGACTTATTTGGTATGTTGTATATATTACATAAAATAACTTTT
CAAATATAGTTTAAATAACACTTAGAAGTGTTTACTTACCTGGAAAATAATTGCTATGCCGTA
CATTCAGAGTGCCCCCTCCCCTGCAAGGCCTTGCCATGATTAACAAGTAAGTTGTTAGTCTT
ACAGATAATTCATGCATTAACAGTTTAAAGATTTAGACCATGGTAATAGTAGTTCTTATTCTC
TAAGGTTATATCATATGTAATTTAAAAGTATTTTAAAGACAAGTTTCCTGTATACCTCTGAA
CTGTTTTGATTTTGAGTTCATCATGATAGATCTGCTGTTTCCTTATAAAAGGCATTTGTTGT
GTGAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACATACCTGACCAA
AAAATTCCCAGTAACCAAGGCATGATCAATTTATAGTGGTCGTTTACATCTAATAATTATCAG
GACTTTTTTCAGGAGTGGGTTATAAAAAACATTCAAGTTGGTCTGACAGTATTTTGTTAAGGA
TATTTGTTGTATGTTTATTCAGTATACTTACATAAAAAATTATTTGCCCATCAGCCAAAAC
CAGTAATCATGACAGCTGTCTGTTGTTTTATGAAGTTTATTTCTCAAGAAAATGGGAATAAA
TTTGGGATTTGTTGAGCTTTTTTACTAAAGATGCCTAAAGCCACAGGTTTTATTGCCTAACT
TAAGCCATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCGGCGTGTGGCTGG
AGCCTTCCCACTGGAGGCTGAAAGTGGCTTGTGGTATTATAATGTTTCAAGAGGAA
GGTGCAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAAGTCTTTGTGCTTGTGAT
CTACTGGACTTTTTTTTTGCAAGGAAGTGCACTCTGCTCCTTCCCTATTTTCTGTTCTGGA
TGTCAGTGCACTGCACTGCTACTGTTTTATCCACTTGGCCACAGACTTTTTCTAACAGCTGC
GTATTATTTCTATATACTAATTGCATTGGCAGCATTGTGTCTTTGACCTGTATACTAGCTT
GACATAGTGCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAATTTTCCATAGAATAT
GCACTGATACAACATTACCATTCTTCTATGGAAAGAAAACTTTTGATGATGAAACAATAAG
ATTTTAAATATCTATTTTAAAAA

—197/310

FIGURE 194

MAGIKALISLSFGGAIGLMFLMLGCALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTDAM
SNACKELAIFLTTGIVVSAFGLPIVFARAHLEWGACALVLTGNTVIFATILGFFLVFGSND
DFSWQQW

[illegible]

199/310

FIGURE 195B

GTAAGTAATCAGAGGGGCAAATGCCACTTGTTATTCCTCCCAAGTTTTCCAAGCAAGTACAC
ACAGATCTCTGGTAGGATTAGGGGCCACTTGTGTTTCCGGCTTATTTTAGTCGACTTGTCAG
CAAGTTTGATGCCTAGTCTATCTGACATGGCCCAGTAGAACAGGGCATTGATGGATCACATG
AGATGGTAGAAGGAACATCATCACATACCCCTCTCACAGAGAAAATTATCAAAGAACCAGAA
ATTATATCTGTTTTGGAGCAAGAGTGTCAATAATGTTTCAGGGTAGTCAAAATAAACATAAAT
TATCTCCTCTAGATGAGTGGCGATGTTGGCTGATTTGGGTCTGCCATTGACAGAATGTCAAA
TAAAAGGAATTAGCTAGAATATGACCATTAAATGTGCTTCTGAAATATATTTTGAGATAGG
TTTAGAATGTCA

200/310

FIGURE 196

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPH
NLSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNNHICSVQGDAFQKLRRVKELTLSSNQ
ITQLPNTTFRPMPNLRSDLSYNKLQALAPDLFHGLRKLTTLHMRANAIQFVPVRIFQDCRS
LKFLDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRILSLHSLCLRRNKVAIV
VSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRITYIEPRIILNSWKSLSIT
LAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDVYAFHLCEDGAEPTSG
HLLSAVTNRSDLGPPASSATTADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMA
LIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHOMAAMSAQEYYVDYKPNH
IEGALVIINEYGSCTCHQQPARECEV

201/310

FIGURE 197

GTGCAAGGAGCCGAGGCGAGATGGGCGTCCTGGGCCGGGTCTGCTGTGGCTGCAGCTCTGC
GCACTGACCCAGGCGGTCTCCAACTCTGGGTCCCCAACACGGACTTCGACGTCGCAGCCAA
CTGGAGCCAGAACCGGACCCCGTGCGCCGGCGGCCCGTTGAGTTCCCGGCGGACAAGATGG
TGTCAGTCCTGGTGCAAGAAGGTCACGCCGTCTCAGACATGCTCCTGCCGCTGGATGGGGAA
CTCGTCCTGGCTTCAGGAGCCGGATTGGGCGTCTCAGACGTGGGCTCGCACCTGGACTGTGG
CGCGGGCGAACCTGCCGTCTTCGCGACTCTGACCGCTTCTCCTGGCATGACCCGCACCTGT
GGCGCTCTGGGGACGAGGCACCTGGCCTCTTCTTCGTGGACGCCGAGCGCGTGCCCTGCCGC
CACGACGACGTCTTCTTCCGCCTAGTGCCTCCTTCGCGTGGGGCTCGGCCCTGGCGCTAG
CCCCGTGCGTGTCGCGAGCATCTCGGCTCTGGGCCGGACGTTACGCGCGACGAGGACCTGG
CTGTTTTCTGCGGTCCCGCGCGGGCCGCCTACGCTTCCACGGGCCGGGCGCGCTAGCGTG
GGCCCCGAGGACTGCGCGGACCCGTGCGGCTGCGTCTGCGGCAACGCGGAGGCGCAGCCGTG
GATCTGCGCGGCCCTGCTCCAGCCCCT

FIGURE 198

MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLVQE
GHAUSDMLLPDGLVLAAGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDPHLWRSGDEA
PGLFFVDAERVPCRHDDVFFPPSASFVGLGPGASPVVRVRSISALGRTFTRDEDLAVFLASR
AGRLRFHGPGLSVGPEDCADPSGCVCGNAEAQPWICAALLQP

FIGURE 200

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIP
FARDAVKKCFVCLA

FIGURE 201

TTGAGCGCAGGTGAGCTCCTGCGCGTTCCGGGGGCGTTCCTCCAGTCACCCTCCCGCCGTTA
CCCCGCGCGCGCCCGAGGGAGTCTCCTCCAGACCCTCCCTCCCGTTGCTCCAACTAATACG
GACTGAACGGATCGCTGCGAGGGTGGGAGAGAAAATTAGGGGGAGAAAGGACAGAGAGAGCA
ACTACCATCCATAGCCAGATAGATTATCTTACACTGAACTGATCAAGTACTTTGAAAATGAC
TTCGAAATTTATCTTGGTGTCTTCATACTTGCTGCACTGAGTCTTTCAACCACCTTTTCTC
TCCAAC TAGACCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACTTA
TATAAAGTTCCAACGCCCCATTTT CATTATATTATGAAATATGGTGTTCACGTGAAGCAAGT
TACTAATGTTTTTATTACAAAAACCTACCCTAACCATTATACTTTGGTAACTGGCCTCTTTG
CAGAGAATCATGGGATTGTTGCAAATGATATGTTTGATCCTATTCCGGAACAAATCTTTCTCC
TTGGATCACATGAATATTTATGATTCCAAGTTTGGGAAGAAGCGACACCAATATGGATCAC
AAACCAGAGGGCAGGACATACTAGTGGTGCAGCCATGTGGCCCGGAACAGATGTAAAAATAC
ATAAGCGCTTTCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAGAGTT
GCCAAAATTGTTGAATGGTTTACGTCAAAAGAGCCCATAAATCTTGGTCTTCTCTATTGGGA
AGACCCTGATGACATGGGCCACCATTGTTGGGACCTGACAGTCCGCTCATGGGGCCTGTCATTT
CAGATATTGACAAGAAGTTAGGATATCTCATACAAATGCTGAAAAAGGCAAAGTTGTGGAAC
ACTCTGAACCTAATCATCACAGTGATCATGGAATGACGCAGTGCTCTGAGGAAAGGTTAAT
AGAAGTTGACCAGTACCTGGATAAAGACCACTATACCCTGATTGATCAATCTCCAGTAGCAG
CCATCTTGCCAAAAGAAGGTAAATTTGATGAAGTCTATGAAGCACTAACTCACGCTCATCCT
AATCTTACTGTTTACAAAAAAGAAGACGTTCCAGAAAGGTGGCATTACAAATACAACAGTCG
AATTCAACCAATCATAGCAGTGGCTGATGAAGGGTGGCACATTTTACAGAATAAGTCAGATG
ACTTCTGTTAGGCAACCACGGTTACGATAATGCGTTAGCAGATATGCATCCAATATTTTTTA
GCCCCATGGTCTCTGCCTTCAGAAAGAATTTCTCAAAAGAAGCCATGAACTCCACAGATTTGTA
CCCCTACTATGCCACCTCCTCAATATCACTGCCATGCCACACAATGGATCATTCTGGAATG
TCCAGGATCTGCTCAATTGAGCAATGCCAAGGGTGGTCCCTTATACACAGAGTACTATACTC
CTCCCTGGTAGTGTTAAACCAGCAGAATATGACCAAGAGGGGTCATACCCTTATTTTCATAGG
GGTCTCTCTTGGCAGCATTATAGTGATTGTATTTTTTGTAAATTTTCATTAAGCATTAAATTC
ACAGTCAAAATACCTGCCTTACAAGATATGCATGCTGAAATAGCTCAACCATTATTACAAGCC
TAATGTTACTTTGAAGTGGATTTGCATATTGAAGTGGAGATTCCATAATTATGTCAGTGT
AAAGGTTTCAAATCTGGGAAACCAGTTCCAAACATCTGCAGAAACCATTAAAGCAGTTACAT
ATTTAGGTATACACACACACACACACACACATACACACACACCGACCAAATACTTACAC
CTGCAAAGGAATAAAGATGTGAGAGTATGTCTCCATTGTTCACTGTAGCATAGGGATAGATA
AGATCCTGCTTTATTTGGACTTGGCGCAGATAATGTATATATTTAGCAACTTTGCACTATGT
AAAGTACCTTATATATTGCACTTTAAATTTCTCTCCTGATGGGTACTTTAATTTGAAATGCA
CTTTATGGACAGTTATGTCTTATAACTTGATTGAAAATGACAACTTTTGCACCCATGTCAC
AGAATACTTGTTACGCATTGTTCAAAGTGAAGGAAATTTCTAATAATCCCGAATAATGAACA
TAGAAATCTATCTCCATAAATTGAGAGAAGAAGAAGGTGATAAGTGTTGAAAATTAATGTG
ATAACCTTTGAACCTTGAATTTTGGAGATGTATCCCAACAGCAGAATGCAACTGTGGGCAT
TTCTTGCTCTTATTTCTTTCCAGAGAACGTGGTTTTTCATTTATTTTTCCCTCAAAAGAGAGTC
AAATACTGACAGATTCGTTCTAAATATATTGTTTTCTGTCATAAAATTATTGTGATTTCTGTA
TGAGTCATATTACTGTGATTTTCTAATAATGAAGACACCATGAATATACTTTTCTTCTATA
TAGTTCAGCAATGGCCTGAATAGAAGCAACCAGGCACCATCTCAGCAATGTTTTCTCTTGTT
TGTAATTATTTGCTCCTTTGAAAATTAATCACTATTAATTACATTAATAATCAAAATGGAT
AAAAAAAAAAAAAAAAAAAA

FIGURE 202

MTSKFILVSFILAALSSTTFSLQLDQQKVLLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK
QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDP IRNKSFSLDHMNIYDSKFWEEATPIW
ITNQRAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY
WEDPDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEER
LIELDQYLDKDHYTELIDQSPVAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN
SRIQPIIAVADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD
LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPYF
IGVSLGSIIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

207/310

FIGURE 203

GGATTTTGTGATCCGCGATTTCGCTCCACGGGCGGGACCTTTGTAAGTGCAGGGAGGCCCAG
GACAGGCCCCACCCTGCGGGGCGGGAGGCAGCCGGGGTGAGGGAGGTGAAGAAACCAAGACGC
AGAGAGGCCAAGCCCCCTTGCCTTGGGTACACAGCCAAAGGAGGCAGAGCCAGAACTCACAA
CCAGATCCAGAGGCAACAGGGACATGGCCACCTGGGACGAAAAGGCAGTCACCCGCAGGGCC
AAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCGTGGGAGACGA
CTACCATGCCTGGAACATCAACTACAAGAAATGGGAGAATGAAGAGGAGGAGGAGGAGGAGG
AGCAGCCACCACCCACACCAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCCTGACGTTGCC
CCTGCCCCCTGGCCCCGCACCCAGGGCCCCCCTTGACTTCAGGGGCATGTTGAGGAACTGTT
CAGCTCCACAGGTTTCAGGTCATCATCATCTGCTTGGTGGTTCTGGATGCCCTCCTGGTGC
TTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCCACAAGAATAACTATGCTGCCATG
GTATTCCACTACATGAGCATCACCATCTTGGTCTTTTTTATGATGGAGATCATCTTTAAATT
ATTTGTCTTCCGCCTGAGTTCTTTCACCACAAGTTTGAGATCCTGGATGCCCGTCGTGGTGG
TGGTCTCATTCATCCTGGACATTGTCCTCCTGTTCCAGGAGCACCAGTTTGAGGCTCTGGGC
CTGCTGATTCTGCTCCGGCTGTGGCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGT
TAAGACACGTTCAGAACGGCAACTCTTAAGGTTAAACAGATGAATGTACAATTGGCCGCCA
AGATTCAACACCTTGAGTTCAGCTGCTCTGAGAAGCCCCCTGGACTGATGAGTTTGCTGTATC
AACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTACTCT
CACACAGCCACCGTGAAAGTCTTGAGTAAATGTGCTGTGTACAGAAGAGAGAGAAGGAAG
CAGGCTGGCATGTTCACTGGGCTGGTGTACGACAGAGAACCTGACAGTCACTGGCCAGTTA
TCACTTCAGATTACAAATCACACAGAGCATCTGCCTGTTTTCAATCACAAGAGAACAAAACC
AAAATCTATAAAGATATTCTGAAAATATGACAGAATTGACAAATAAAAGCATAAACGTGTA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

208/310

FIGURE 204

MATWDEKAVTRRAKVAPAERMSKFLRHFTVVGDDYHAWNINYKKWENEEEEEEEEQPPPTPV
SGEEGRAAAPDVAPAPGPAPRAPLDFRGMLRKLFSSHRFQVIIICLVVLDALLVLAELILD
KIIQPDKNYAAMVFHYMSITILVFFMMEIIFKLFVFRLLSSFTTSLRSWMPVVVVVSFILD
VLLFQEHQFEALGLLILLRLWRVARIINGIIISVKTRSERQLLRLKQMNVLAAKIQHLEFS
CSEKPLD

-208/310

FIGURE 205

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCACCAGCAGGCCGCCAACAT
GCTCTGTCTGTGCCTGTACGTGCCGGTTCATCGGGGAAGCCCAGACCGAGTTCCAGTACTTTG
AGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCCATTTTCAAGCTCAGTGTCTTCATCCCCTCC
CAGGAATTCTCCACCTACCGCCAGTGGAGCAGAAAAATTGTACAAGCTGGAGATAAGGACCT
TGATGGGCAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGA
GGCTGGTGTTTAAGATTTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATG
CAGTCCCTGCCGGGACTTGGGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCTCAAGAG
CATGGATAAAAAACGGCAGCATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCC
ACCCCGTGGAACATCCCGAGATCATCCTCTACTGGAAGCATTCCACGATCTTTGATGTG
GGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAGAGGCAGACGGGGATGTGGTG
GAGACACCTGGTGGCAGGAGGTGGGGCAGGGGCGGTATCCAGAACCTGCACGGCCCCCTGG
ACAGGCTCAAGGTGCTCATGCAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGT
GGCTTCACTCAGATGATTTCAGAGAAGGAGGGGCCAGGTCACTCTGGCGGGGCAATGGCATCAA
CGTCCTCAAAATTGCCCCGAATCAGCCATCAAATTCATGGCCTATGAGCAGATCAAGCGCC
TTGTTGGTAGTGACCAGGAGACTCTGAGGATTACAGAGAGGCTTGTGGCAGGGTCCTTGGCA
GGGGCCATCGCCAGAGCAGCATCTACCCAATGGAGGTCTGAAGACCCGGATGGCGCTGCG
GAAGACAGGCCAGTACTCAGGAATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAGGGGG
TGGCCGCTTCTACAAAGGCTATGTCCCCAACATGCTGGGCATCATCCCCTATGCCGGCATC
GACCTTCGCAGTCTACGAGACGCTCAAGAATGCTGGTGGCTGCAGCACTATGCAGTGAACAGCGC
GGACCCCGCGTGTGTGCTCCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGG
CCAGCTACCCCTGGCCCTAGTCAAGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCT
CCGGAGGTGACCATGAGCAGCCTCTTCAAACATATCCTGCGGACCGAGGGGGCCTTCGGGCT
GTACAGGGGGCTGGCCCCCACTTCATGAAGGTATCCAGCTGTGAGCATCAGCTACGTGG
TCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTTCGCGGTGACGGGGGGAGGGCCGCCCG
GCAGTGGACTCGCTGATCCTGGGCGCAGCCTGGGGTGTGCAGCCATCTCATTCTGTGAATG
TGCCAACTAAGCTGTCTCGAGCCAAGCTGTGAAAACCTAGACGCACCCGCAGGGAGGGT
GGGGAGAGCTGGCAGGCCCAGGGCTTGTCTGCTGACCCAGCAGACCCCTCCTGTTGGTTCC
AGCGAAGACCACAGGCATTCTTAGGGTCCAGGGTCAGCAGGCTCCGGGCTCACATGTGTAA
GGACAGGACATTTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGGCTTAGT
TCTTCCATTTACCCCTTGCGAGCCAGCTGTTGGCCACGGCCCCCTGCCCTCTGGTCTGCCGTGC
ATCTCCCTGTGCCCTCTTGCTGCCTGCCTGTCTGCTGAGGTAAGGTGGGAGGAGGGCTACAG
CCCACATCCCACCCCTCGTCCAATCCCATATCCATGATGAAAGGTGAGGTACGTGGCCT
CCCAGGCCTGACTTCCCAACCTACAGCATTGACGCCAACTTGGCTGTGAAGGAAGAGGAAAG
GATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATCT
TGGGAGTGCAGGGGGCTCGGGCTGCCTGGCCTGGCTGCACAGAAGGCAAGTGTCTGGGGCTCA
TGGTGTCTGAGCTGGCCTGGACCCTGTCAGGATGGGCCCCACCTCAGAACCAACTCACTG
TCCCCACTGTGGCATGAGGGCAGTGGAGCACCATGTTTGAGGGCGAAGGGCAGAGCGTTTGT
GTGTTCTGGGGAGGGAAGGAAAAGGTGTTGGAGGCCCTTAATTATGGAAGTGTGGGAAAAGGG
TTTTGTCCAGAAGGACAAGCCGGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCCTGTCCAACC
CCAGCAGGGGCGCAGCGGGACCAGCCCCACATTCCACTTGTGTCACTGCTTGGAACTTATTT
ATTTTGTATTTATTTGAACAGAGTTATGTCTTAATTTTATAGATTTGTTTAAATTAATA
GCTTGTCATTTTCAAGTTCATTTTTTATTCATATTTATGTTTCATGTTTGAATTGTACCTTCCC
AAGCCCCCCCAGTGGGATGGGAGGAGGAGGAGAAGGGGGGCTTGGGCCGCTGCAGTCACAT
CTGTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAAGGCAGCAGCCCTG
GCTCCTTTCTTTTGGCAGGTTGGGGAAGGGCTTGCCCCAGCCTTAGGATTTTCAAGGTTTGA
CTGGGGGCGTGGAGAGAGAGGGAGGAACCTCAATAACCTTGAAGGTGGAATCCAGTTATTTT
CTGCGTGCAGGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAGTGAGGTGCCTCT
CAGCTGAAATTTGTGGTGGGCGGGGCTGGAGGAGGGTGGGGGCTGGCTTCCCTCCCTCC
CAGCCTTCTGCTGCCCTTGCTTAACAATGCCGGCCAACCTGGCGACCTCACGGTTCAGCTTCC
ATTCCACCAGAATGACCTGATGAGGAAATCTTCAATAGGATGCAAAGATCAATGCAAAAATT
GTTATATATGAACATATACTGGAGTCTGCAAAAAGCAAATTAAGAAAGAATTGGACGTTAG
AAGTTGTCAATTTAAAGCAGCCTTCTAATAAAGTTGTTTCAAAGCTGAAAAAAAAAAAAAA
AA

FIGURE 206

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKD
LDGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEEKILK
SMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMW
WRHLVAGGGAGAVSRTCTAPLDRLKVLMOVHASRSNNMGIVGGFTQMIREGGARSLWRGNGI
NVLKIAPESAIKFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQSSIYPMEVLKTRMAL
RKTGQYSGMLDCARRILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNS
ADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFG
LYRGLAPNFMKVIPAVSISYVVYENLKITLGVQSR

212/310

FIGURE 208

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDI
KLSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASRLKKNVQLTD
AGTYKCYIIITSKGKGANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVWVASQVD
QGANFSEVSNTSFELNSENVTMKVVSVLNVNTINNTYSCMIENDIAKATGDIKVTSEIKRR
SHLQLLNSKASLCVSSFFAISWALLPLSPYMLK

213 / 310

FIGURE 209

[illegible]

214/310

FIGURE 210

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLQEMKTLFLNTEYLMPFLLNQCSSLLYLTLASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGSRHTCVSSFPEPISPEWVRTRPFPILPFPLQLFCFLVAIRVPFPWTVWRKTEAGVWD

215/310

FIGURE 211

CTTCTGTAGGACAGTCACCAGGCCAGATCCAGAAGCCTCTCTAGGCTCCAGCTTTCTCTGTG
GAAGATGACAGCAATTATAGCAGGACCCTGCCAGGCTGTGAAAAGATTCCGCAATAAACT
TTGCCAGTGGGAAGTACCTAGTGAAACGGCCTAAGATGCCACTTCTTCTCATGTCCCAGGCT
TGAGGCCCTGTGGTCCCCATCCTTGGGAGAAGTCAGCTCCAGCACCATGAAGGGCATCCTCG
TTGCTGGTATCACTGCAGTGCTTGTTCAGCTGTAGAATCTCTGAGCTGCGTGCAAGTGAAT
TCATGGGAAAAATCCTGTGTCAACAGCATTGCCTCTGAATGTCCCTCACATGCCAACACCAG
CTGTATCAGCTCCTCAGCCAGCTCCTCTCTAGAGACACCAGTCAGATTATACCAGAATATGT
TCTGCTCAGCGGAGAAGTGCAGTGAGGAGACACACATTACAGCCTTCACTGTCCACGTGTCT
GCTGAAGAACAACCTTTTCAATTTTGTAAAGCCAGTGCTGCCAAGGAAAGGAATGCAGCAACACCAG
CGATGCCCTGGACCCTCCCCTGAAGAACGTGTCCAGCAACGCAGAGTGCCCTGCTTGTTATG
AATCTAATGGAACTTCCTGTGCTGGGAAGCCCTGGAAATGCTATGAAGAAGAACAGTGTGTC
TTTCTAGTTGCAGAACTTAAGAATGACATTGAGTCTAAGAGTCTCGTGCTGAAAGGCTGTTC
CAACGTCAGTAACGCCACCTGTCAGTTCCTGTCTGGTGAAAACAAGACTCTTGGAGGAGTCA
TCTTTTCGAAAGTTTGAGTGTGCAAATGTAAACAGCTTAACCCCCACGTCTGCACCAACCACT
TCCCACAACGTGGGCTCCAAAGCTTCCCTCTACCTCTTGGCCCTTGCCAGCCTCCTTCTTCG
GGGACTGCTGCCCTGAGGTCTTGGGGCTGCACTTTGCCCAGCACCCCATTTCTGCTTCTCTG
AGGTCCAGAGCACCCCCTGCGGTGCTGACACCCTCTTTCCCTGCTCTGCCCCGTTTAACTGC
CCAGTAAGTGGGAGTCACAGGTCTCCAGGCAATGCCGACAGCTGCCTTGTTCTTCATTATTA
AAGCACTGGTTCATTCACTGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

215/310

FIGURE 212

MKGILVAGITAVLVAAVESLSCVQCNSWEKSCVNSIASECPSHANTSCISSSASSSLETPVR
LYQNMFCSAENCSEETHITAFTVHVSAEEHFHFVSQCCQGKECSNTSDALDPPLKNVSSNAE
CPACYESNGTSCRGKPWKCYEEEQCVFLVAELKNDIESKSLVLKGCSNVSNATCQFLSGENK
TLGGVIFRKFECANVNSLTPTSAPTTSNHNVGSKASLYLLALASLLLRGLLP

217/310

FIGURE 213

GGCCTCGGTTCAAACGACCCGGTGGGTCTACAGCGGAAGGGAGGGAGCGAAGGTAGGAGGCA
GGGCTTGCCTCACTGGCCACCCTCCCAACCCCAAGAGCCCAGCCCCATGGTCCCCGCCGCCG
GCGCGCTGCTGTGGGTCTGCTGCTGAATCTGGGTCCCCGGGCGGCGGGGGCCCAAGGCCTG
ACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGCTTTGGGGGCCCCATGACCCGCAG
CTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTAGAGGACGAGA
ATGATGCCATGGCCGACGCCGACCGCCTGGCTGGACCAGCGGCTGCCGAGCTCTTGGCCGCC
ACGGTGTCCACCGGCTTTAGCCGGTCGTCCGCCATTAACGAGGAGGATGGGTCTTCAGAAGA
GGGGGTTGTGATTAATGCCGGAAGGATAGCACCAGCAGAGAGCTTCCCAGTGCGACTCCCA
ATACAGCGGGGAGTTCCAGCACGAGGTTTATAGCCAATAGTCAGGAGCCTGAAATCAGGCTG
ACTTCAAGCCTGCCGCGCTCCCCGGGAGGTCTACTGAGGACCTGCCAGGCTCGCAGGCCAC
CCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTCACCCCTCACCCACAG
CCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCAC
TGCAAGTCGGGCACCATGAGCCGAGCCGGTCTGGGAAGCTGCACGGCCTTTCGGGGCGCCT
TCGAGTTGGGGCGCTGAGCCAGCTCCGCACGGAGCACAAGCCTTGCACCTATCAACAATGTC
CCTGCAACCGACTTCGGGAAGAGTGCCCCCTGGACACAAGTCTCTGTACTGACACCAACTGT
GCCTCTCAGAGCACCAACAGTACCAGGACCACCACTACCCCCTTCCCCACCATCCACCTCAG
AAGCAGTCCCAGCCTGCCACCCGCCAGCCCCTGCCAGCCCTGGCTTTTTGGAAACGGGTCA
GGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTACAGAGATGCAACCAATA
GACAGAAACCAGAGGTAATGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCT
CTTGCCCTTTCAATCCTAGCACCCACTAGATATTTTGTAGTACAGAAAAACAAACTGGAAAA
CACAA

218/310

FIGURE 214

MVPAAGALLWVLLLNLGPRAAGAQGLTQTPTEMQRVSLRFGGPMTRSYRSTARTGLPRKTRI
ILEDENDAMADADRLAGPAAAEELLAATVSTGFSRSSAINEEDGSSEEGVVINAGKDSTSREL
PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP
SPSPTAMPSPEDLRLVLMWPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC
TYQQCPCNRLREECPLDTSLCTDTNCASQSTTSTRTTTTPFPTIHLRSSPSLPPASPCPALA
FWKRVIRIGLEDIWNSLSSVFTEMQPIDRNQR

219/310

FIGURE 215

CCCGGGTTCGACCCACGCGTCCGGGGAGAAAGGATGGCCGGCCTGGCGGCGCGGTTGGTCCTG
CTAGCTGGGGCAGCGGCGCTGGCGAGCGGCTCCCAGGGCGACCGTGAGCCGGTGTACCGCGA
CTGCGTACTGCAGTGCAGAGAGCAGAACTGCTCTGGGGGCGCTCTGAATCACTTCCGCTCCC
GCCAGCCAATCTACATGAGTCTAGCAGGCTGGACCTGTGCGGACGACTGTAAGTATGAGTGT
ATGTGGGTACACGTTGGGCTCTACCTCCAGGAAGGTCAAAAGTGCCTCAGTTCCATGGCAA
GTGGCCCTTCTCCCGGTTCTGTCTTTCAAGAGCCGGCATCGGCCGTGGCCTCGTTTCTCA
ATGGCCTGGCCAGCCTGGTGATGCTCTGCCGCTACCGCACCTTCGTGCCAGCCTCCTCCCCC
ATGTACCACACCTGTGTGGCCTTCGCCTGGGTGTCCCTCAATGCATGGTTCTGGTCCACAGT
CTTCCACACCAGGGACACTGACCTCACAGAGAAAATGGACTACTTCTGTGCCTCCACTGTCA
TCCTAACTCAATCTACCTGTGCTGCGTCAGGACCGTGGGGCTGCAGCACCCAGCTGTGGTC
AGTGCCTTCCGGGCTCTCCTGCTGCTCATGCTGACCGTGCACGTCTCCTACCTGAGCCTCAT
CCGCTTCGACTATGGCTACAACCTGGTGGCCAACGTGGCTATTGGCCTGGTCAACGTGGTGT
GGTGGCTGGCCTGGTGCTGTGGAACCAGCGGCGGCTGCCTCACGTGCGCAAGTGCCTGGTG
GTGGTCTTGCTGCTGCAGGGGCTGTCCCTGCTCGAGCTGCTTGACTTCCACCGCTCTTCTG
GGTCTGGATGCCCATGCCATCTGGCACATCAGCACCATCCCTGTCCACGTCCTCTTTTCA
GCTTCTGGAAGATGACAGCCTGTACCTGCTGAAGGAATCAGAGGACAAGTTCAAGCTGGAC
TGAAGACCTTGGAGCGAGTCTGCCCCAGTGGGGATCCTGCCCCCGCCCTGCTGGCCTCCCTT
CTCCCCTCAACCCTTGAGATGATTTTCTCTTTTCAACTTCTTGAACCTTGGACATGAAGGATG
TGGGCCCAGAATCATGTGGCCAGCCACCCCTGTGGCCCTCACAGCCTTGGAGTCTGTT
CTAGGGAAGGCCTCCAGCATCTGGGACTCGAGAGTGGGCAGCCCTCTACCTCCTGGAGCT
GAACTGGGGTGGAACTGAGTGTGTTCTTAGCTCTACCGGGAGGACAGCTGCCTGTTTCTCC
CCACCAGCCTCCTCCCCACATCCCCAGCTGCCTGGCTGGGTCTGAAGCCCTCTGTCTACCT
GGGAGACCAGGGACCACAGGCCTTAGGGATACAGGGGGTCCCCCTCTGTTACCACCCCCAC
CCTCCTCCAGGACACCACTAGGTGGTGCTGGATGCTTGTTCTTTGGCCAGCCAAGGTTACAG
GCGATTCTCCCCATGGGATCTTGAGGGACCAAGCTGCTGGGATTGGGAAGGAGTTTCACCCT
GACCGTTGCCCTAGCCAGGTTCACAGGAGGCCTCACCATACTCCCTTTCAGGGCCAGGGCTC
CAGCAAGCCCAGGGCAAGGATCCTGTGCTGCTGTCTGGTTGAGAGCCTGCCACCGTGTGTG
GGAGTGTGGGCCAGGCTGAGTGCATAGGTGACAGGGCCGTGAGCATGGGCCTGGGTGTGTGT
GAGCTCAGGCCTAGGTGCGCAGTGTGGAGACGGGTGTTGTGCGGGAAGAGGTGTGGCTTCAA
AGTGTGTGTGTGCAGGGGGTGGGTGTGTTAGCGTGGGTTAGGGGAACGTGTGTGCGCGTGCT
GGTGGGCATGTGAGATGAGTGAAGTGCAGGGCCGTGAGCATGGGCCTGGGTGTGTGT
ATGAGGGAATCCTGTCAACATCAATAACTACTTGTGGAGCGCCAGCTCTGCCCAGACGCCA
CCTGGGCGGACAGCCAGGAGCTCTCCATGGCCAGGCTGCCTGTGTGCATGTTCCCTGTCTGG
TGCCCCCTTTGCCCGCCTCCTGCAAACCTCACAGGGTCCCCACACAACAGTGCCCTCCAGAAG
CAGCCCCCTCGGAGGCAGAGGAAGGAAAATGGGGATGGCTGGGGCTCTCTCCATCCTCCTTTT
CTCCTTGCCCTTCGCATGGCTGGCCTTCCCCCTCCAAAACCTCCATTCCCCTGCTGCCAGCCCC
TTTGCCATAGCCTGATTTTGGGGAGGAGGAAGGGGCGATTTGAGGGAGAAGGGGAGAAAGCT
TATGGCTGGGTCTGGTTTCTTCCCTTCCCAGAGGGTCTTACTGTTCCAGGGTGGCCCCAGGG
CAGGCAGGGGCCACACTATGCCTGTGCCCTGGTAAAGGTGACCCCTGCCATTTACCAGCAGC
CCTGGCATGTTCTGCCCCACAGGAATAGAATGGAGGGAGCTCCAGAACTTTCCATCCCAA
AGGCAGTCTCCGTGGTTGAAGCAGACTGGATTTTGTCTGCCCCCTGACCCCTTGTCCCTCT
TTGAGGGAGGGGAGCTATGCTAGGACTCCAACCTCAGGGACTCGGGTGGCCTGCGCTAGCTT
CTTTTGATACTGAAAACCTTTAAGGTGGGAGGGTGGCAAGGGATGTGCTTAATAAATCAATT
CCAAGCCTCAAAAAAAAAAAAAAAAAA

-e20/310

FIGURE 216

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMSLAGW
TCRDDCKYECMWVTVGLYLQEGHKVPQFHGKWPF SRFLFFQEPASAVASFLNGLASLVMLCR
YRTFVPASSPMYHTCVAFVSLNAFWSTVFHTRDLDLTKMDYFCASTVILHSIYLCCVR
TVGLQHPAVVSAFRALLLMLTVHVSYSLSLIRFDYGYNLVANVAIGLVNVVWWLAWCLWNQR
RLPHVRKCVVVVLLLQGLSLELLDFPPLFWVLDAAHAIWHISTIPVHVLFFSFLEDDSLYLL
KESEDKFKLD

FIGURE 217

[illegible]

222/310

FIGURE 218

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPEKSSTKETERKETKAEEL
DAEVLEVFPHTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTN
TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV
RLINKFNSSSSSLEEKIAALFDLEYVYVHQMDNAQDLLSFGGLQVVINGLNSTEPLVKEYAAF
VLGAAFSSNPQVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL
GGLQVLRTLVQEKGTEVLAVRVVTLVTEKMFEEEEAEELTQEMSPEKLQYRQVHLLPG
LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLAS
LELQDGEDEGYFQELLGSVNSLLKELR

[illegible]

FIGURE 220

MGAAVFVGCTFVAFGPAPALFLITVAGDPLRVIIILVAGAFFWLVSLLLASVVWFILVHVTDR
SDARLQYGLLIIFGAAVSVLLQEVFRFAYYKLLKKADEGLASLSEDGRSPISIRQMAYVSGLS
FGIISGVFSVINILADALGPGVVGIIHGDSPIYFLTSAFLTAIIILLHTFWGVVFFDACERRR
YWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMGLWAFITAGGSLRSIQRSLCKD

225/310

FIGURE 221

AAGCTGGTTTAAGGAAGCAGAGGAGGGTTAGATTCGTTGAGTGAGGACGGAAGATCAACCCA
TTTCCATTCCGCCAGATGGCCTATGTTTCTGGTCTCTCCCTTCGGNATCATCAGTGGTGTNT
TNTCTGTTATCAATATTTTGGCTGATGCANTTGGGCCAGGTGTGGTTGGGATCCATGGAGAC
TCACCCTATTANTTCCTGANTTCAGCCTTTNTGACAGCAGCCATTATCCTGCTC

225/310

FIGURE 222

GACCGACCGTTCAGATGCCCCGGTTCAGTACGGCTTCCTGATTTTGGTGCTGCTGTNTCTG
TCCTTCTACAGGAGGTGTTCCGCTTTGCCTANTACAAGCTGCTTAAGAAGGCAGATGAGGGG
TTAGCATNGCTGAGTGAGGACGGAAGATCACCCATTTCCATCCGCCAGATGGCCTATGTTTN
TGGTNTTTCCTTCGGTATCATCAGTGGTGTNTNTCTGTTATCAATATTTGGNTGATGCAN
TTGGGCCAGGTGTGGTTGGGATCCATGGAGANTCACCCATTAAATTCCTGAATTCAGCCTTT
NTGACAGCAGCCATTATCCTGNTCCATACCTTTTGGGGAGTTGTGTTTTTTGATGCCTGTGA
GAGGAG

227/310

FIGURE 223

NGTTGGAGAAGTGGCGCGGACNTTCATTTGGGGTTTCGGTTTCCCCCTTTCCCTTTCCCCG
GGGTCTGGGGTGACATTGCACGGGCCCCCTCGTGGGGTCGCGTTGCCACCCACGCGGACTCC
CCAGNTGGNGCGCCCTTCCCATTTCCTGTCTGGTCAGGCCCCACCCCCCTTCCCACNTG
ACCAGCCATGGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTGCGGTTGCGCCCGGCCTTCG
CGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGCA
TTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGAC
CGACCGGTGAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTCC
TTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTTG
GGCCAGGTGTGGTTGGGATCCATGGAGACTCACCC

-228/310

FIGURE 224

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGGTTCCCCCCTTTCCCNNTTCCCCGGGG
TCTGGGGGTGACATTGCACCGCGCCCNCTCGTGGGGTTCGCGTTGCCACCCACGCGGACTCCC
CAGNTGGCGCGCCCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCTTCCCACCTGA
CCAGCCATGGGGGCTGCGGTGTTTTTCGGGCTGCACTTTCGTGCGGTTTCGGGCCCGGCCTTC
GCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGC
ATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGA
CCGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTTGGTGCTGCTGTCTCTGTC
CTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTT
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG
GTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

ee9/310

FIGURE 225

GCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCGGTGCCCAGAGCCCAGGAGGAGGCAG
TGGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCACCCCC
TACCTGGGGGACAGGGCAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCAGGGAAGGAGAGG
TGTCTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCTCCTTGCCCTGTCTGGAGGCTGCT
AGACTCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCGT
CCTTGTGGTTCCTCTCTACCTGGGGAAATAAGGTGCAGCGGCCATGGCTACAGCAAGACCCC
CCTGGATGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGTACAGAGCATGTT
CTCGCCAACAATGATGTTTCCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCA
GGACCTGGGAGCTGGGGCCGGGGAAGACGCCCCGGTCGGATGACAGCAGCAGCCGCATCATCA
ATGGATCCGACTGCGATATGCACACCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAAC
CAGCTCTACTGCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAG
GAAGAAAGTTTTAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGC
AGCAGATGTTCCAGGGGGTCAAATCCATCCCCACCCTGGCTACTCCCACCCTGGCCACTCT
AACGACCTCATGCTCATCAAAGTGAACAGAAGAATTGTCCTCCACTAAAGATGTCAGACCCAT
CAACGTCTCCTCTCATTGTCCCTCTGCTGGGACAAAGTGCTTGGTGTCTGGCTGGGGGACAA
CCAAGAGCCCCCAAGTGCCTTCCCTAAGGTCTCCAGTGCTTGAATATCAGCGTGCTAAGT
CAGAAAAGGTGCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGA
CAAAGCAGGTAGAGACTCCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCC
TGCAGGGACTCGTGTCTGGGGAGATTACCCTTGTGCCCGGCCAACAGACCGGGTGTCTAC
ACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCCTGAGTCAT
CCCAGGACTCAGCACACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTCAG
ACCCTCATTCCTTCCCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCCTGACCCCATGTCT
CCTGGACTCAGGGTCTGCTTCCCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGG
GAACAATTTCCAAAAGTGTCCAGGGCGGGGGTTCGCTCTCAATCTCCCTGGGGCACTTTCAT
CCTCAAGCTCAGGGCCCCTCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAA
CTGAGAAGTGGAATAAAAAA

-230/310

FIGURE 226

MATARPPWMWVLCALITALLGVTEHVLANNNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDD
SSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLS
PVYESGQQMFQGVKSIPHPGYSHPGHSNDLMLIKLNRRI RPTKDVRPINVSSHCP SAGTKCL
VSGWGTTKSPQVHF PKVLQCLNISVLSQKRCE DAYPRQIDDTMFCAGDKAGR DSCQGDSGGP
VVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

231 / 310

FIGURE 227

ATGGTCAACGACCGGTGGAAGACCATGGGCGGCGCTGCCCAACTTGAGGACCGGCCGCGGA
CAAGCCGACGCGGCCGAGCTGCGGCTACGTGCTGTGCACCGTGCTGCTGGCCCTGGCTGTGC
TGCTGGCTGTAGCTGTACCGGTGCCGTGCTCTTCTGAACACGCCCACGCGCCGGGCACG
GCGCCCCACCTGTGTCGTAGCACTGGGGCTGCCAGCGCCAACAGCGCCCTGGTCACTGTGA
AAGGGCGGACAGCTCGCACCTCAGCATCCTCATTGACCCGCGCTGCCCCGACCTCACCGACA
GCTTCGCACGCCTGGAGAGCGCCAGGCCTCGGTGCTGCAGGCGCTGACAGAGCACCAGGCC
CAGCCACGGCTGGTGGGCGACCAGGAGCAGGAGCTGCTGGACACGCTGGCCGACCAGCTGCC
CCGGCTGCTGGCCCGAGCCTCAGAGCTGCAGACGGAGTGCATGGGGCTGCGGAAGGGGCATG
GCACGCTGGGCCAGGGCCTCAGCGCCCTGCAGAGTGAGCAGGGCCGCTCATCCAGCTTCTC
TCTGAGAGCCAGGGCCACATGGCTCACCTGGTGAACTCCGTACGCGACATCCTGGATGCCCT
GCAGAGGGACCGGGGGCTGGGCCGGCCCCGCAACAAGGCCGACCTTCAGAGAGCGCCTGCC
GGGGAACCCGGCCCCGGGGCTGTGCCACTGGCTCCCGCCCCGAGACTGTCTGGACGTCTCTC
CTAAGCGGACAGCAGGACGATGGCGTCTACTCTGTCTTTCCACCCACTACCCGGCCGGCTT
CCAGGTGTACTGTGACATGCGACGGACGGCGGGCTGGACGGTGTTCAGCGCCGGGAGG
ACGGCTCCGTGAACTTCTTCCGGGGCTGGGACGCGTACCGAGACGGCTTTGGCAGGCTCAC
GGGAGCACTGGCTAGGGCTCAAGAGGATCCACGCCCTGACCACACAGGCTGCCTACGAGCT
GCACGTGGACCTGGAGGACTTTGAGAATGGCACGGCCTATGCCCGCTACGGGAGCTTCGGCG
TGGGCTTGTCTCCGTGGACCCTGAGGAAGACGGGTACCCGCTACCGTGGCTGACTATTCC
GGCACTGCAGGCGACTCCCTCCTGAAGCACAGCGGCATGAGGTTACCACCAAGGACCGTGA
CAGCGACCATTCAGAGAACTGTGCCGCTTCTACCGCGGTGCCTGGTGGTACCGCACT
GCCACACGTCCAACCTCAATGGGCAGTACCTGCGCGGTGCGCACGCCCTCCTATGCCGACGGC
GTGGAGTGGTCCTCCTGGACCGGCTGGCAGTACTCACTCAAGTTCTCTGAGATGAAGATCCG
GCCGTCCGGGAGGACCGCTAGACTGGTGACCTGTCTTGGCCCTGCTGGTCCCTGTGCG
CCCATCCCCGACCCACCTCACTCTTTCGTGAATGTTCTCCACCCACCTGTGCCTGGCGGAC
CCACTCTCCAGTAGGGAGGGGCGGGGCCATCCCTGACACGAAGCTCCCTGGGCCGGTGAAGT
CACACATCGCCTTCTCGCCGTCCCCACCCCTCCATTTGGCAGCTCACTGATCTCTTGCCTC
TGCTGATGGGGCTGGCAAACCTTGACGACCCCACTCCTGCCTGCCCCCACTGTGACTCCGG
TGCTGTTTGGCCTCCCTGGCCAGGATGGTGGAGTCTGCCCCAGGCACCCCTCTGCCCTGCC
GGCCAAATACCCGGCATTATGGGGACAGAGAGCAGGGGGCAGACAGCACCCTGGAGTCCCTC
CTAGCAGATCGTGGGGAATGTGAGGTCTCTCTGAGGTGAGGTCTGAGGCCAGTATCCTCCAG
CCCTCCCAATGCCAACCCCAACCCCGTTTCCCTGGTGCCAGAGAACCACCTCTCCCCAA
GGGCTCAGCCTGGCTGTGGGCTGGGTGGCCCCATCCTACCAGGCCCTGAGGTGAGGATGGG
GAGCTGCTGCCTTTGGGGACCCACGCTCCAAGGCTGAGACCAGTTCCCTGGAGGCCACCCAC
CCTGTGCCCCGGCAGGCCTGGGGTCTGCAGTCTCTTACCTGCTGTGCCACCTGCTCTCTG
TCTCAAATGAGGCCCCAACCCATCCCCACCCAGCTCCCGCCGCTCCTCCTACCTGGGGCAGC
CGGGGCTGCCATCCCATTCTCCTGCCTCTGGAAGGTGGGTGGGGCCCTGCACCGTGGGGCT
GGACTGCGCTAATGGGAAGCTCTTGGTTTTCTGGGCTGGGGCCTAGGCAGGGCTGGGATGAG
GCTTGTAACAACCCCAACCAATTTCCAGGGACTCCAGGGTCTGAGGCCCTCCAGGAGG
GCCTTGGGGGTGATGACCCCTTCCCTGAGGTGGTGTCTCCATGAGGAGGCCAACCTTGCC
ATTGACCGTGGCCACCTGGACCCAGGCCAGGCCCGCCCGCGAGTGGTCAAGGGACAGGGA
CCACCTCACCGGGCAAATGGGGTGGGGGGACTGGGGCACCAGACCAGGCACCACCTGGACA
CTTTCTTGTGAATCCTCCCAACACCCAGCACGCTGTATCCCCACTCCTTGTGTGCACACA
TGCAGAGGTGAGACCCGAGGCTCCCAGGACCAGCAGCCACAAGGGCAGGGCTGGAGCCGGG
TCCTCAGCTGTCTGCTCAGCAGCCCTGGACCCGCGTGCGTTACGTACGGCCAGATGCAGGG
CGGCTTTTCCAAGGCCTCCTGATGGGGGCTCCGAAAGGCTGGAGTCAGCCTTGGGGAGCT
GGCTAGCAGCCTCTCCTCGGGCAGGAGGGGAGGTGGCTTCTCCAAAGGACACCCGATGGCA
GGTGCCTAGGGGGTGTGGGGTTCGTTCTCCCTTCCCTCCCACTGAAGTTTGTGCTTAAAA
AACATAAAATTTGACTTGGCACCCTGGGGGTGGTGGGAGAGGCCGTGTGACCTGGCTCTC
TGTCCAGTGCCACCAGGTCATCCACATGCGCAG

232/310

FIGURE 228

MVNDRWKTMGGAAQLEDRPRDKPQRPSCGYVLCTVLLALAVLLAVAVTGAVLFLNHAHAPGT
APPPVVSTGAASANSALVTVERADSSHLSILIDPRCPDLTDSFARLESAQASVLQALTEHQA
QPRLVGDQEQELDLADQLPRLLARASELQTECMGLRKGHGTLGQGLSALQSEQGRLIQLL
SESQGHMAHLVNSVSDILDALQRDRGLGRPRNKADLQRAPARGTRPRGCATGSRPRDCLDVL
LSGQQDDGVYSVFPTHYPAGFQVYCDMRTDGGGWTVFQRREDGSVNFFRGWDAYRDGFGRLT
GEHWLGLKRIHALTTQAAVELHVDLEDFENGTAARYGSFGVGLFSVDPEEDGYPLTVADYS
GTAGDSLLKHSGMRFTTKDRSDHSENNCAAFYRGAWWYRNCHTSNLNGQYLRGAHASADG
VEWSSWTGWQYSLKFSEMKIRPVREDR

FIGURE 229

GCAGTCAGAGACTTCCCCTGCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCT
TGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCAGGGCAATCCGACCACATTTCACTCT
CACCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGCTGGATGATG
ATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACTCGGCATCCAGAGCCC
CGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGAC
TTTGTGCTTGGTGTCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTACTACC
AGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCC
CAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAAGGAAGTCTGCAGCATGTGGC
TGAAAACTCTGTCTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAG
AACAATGGAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAG
GACTGTAAATATTTCTGCCTTAGTGAAAACCTCTACCATGCTGAAGATAAACAAACAAGAAGA
CCTGGAATTTGCCGCTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTT
TGCGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTTCACCTTCTGAACTG
TTCCATATTATAATAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGG
GATGATCTTCTCAAAGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAA
TGGTGAAGCCAGAGAGCCTCCATGTCCCCCTGAAACATTAGGCCAAGGTGACTGATTGCCC
CTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACAT
TGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCTCGTG
TTTCCTGTTTACAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACA
AGAAGTCTTATTTACATGCCACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTG
GCTTAGAGATAACTTTTAGCTCTCTTTCTTCTCAATGTCTAATATCACCTCCCTGTTTTCAT
GTCTTCCTTACACTTGGTGGAATAAGAACTTTTTGAAGTAGAGGAAATACATTGAGGTAAC
ATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACC
AGCAAATACACAAGGAATTCTTTTTGTTTGTTCAGTTCATACTAGTCCCTTCCCAATCCAT
CAGTAAAGACCCCATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAG
AATCTCAAATCTCAATGCCTTATAAGCATTCCCTTCTGTGTCCATTAAGACTCTGATAATTG
TCTCCCCTCCATAGGAATTTCTCCCAGGAAAGAAATATATCCCCATCTCCGTTTCATATCAG
AACTACCGTCCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAGTGAGCCTCTTCA
TCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAG
GTACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

234/310

FIGURE 230

MQAKYSSTRDMLDDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCLVLL
IGLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRE
LYNKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAAS
QSYSEFFYSYWTGLLRPD SGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKD
CKELKRCVCERRAGMVKPESLHVPPETLGEGD

235/310

FIGURE 231

AATTTTCACCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGNTGG
ATGATGATGGGACACCACCATGAGCCTGCATTNTCAAGCTTTTGCCACAATTCGGCATCCAG
AGCCCCGGCGCACAGAGCACAGGGNTCCTTTTTCAACGTGGCGACCAGTGGCCCTGACCCTG
CTGACTTTGTGCTTGCTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTCAGTA
CTACCAGCTCTCCAATACTGGTCAAGACACCATTCTCAAATGGAAGAAAGATTAGGAAATA
CGTCCCAAGAGTTGCAATTTNTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTNTGCAGCAT
GTGGCTGAAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGGAACCTTTGAAGGAGGGCAA
AGTNTCCTCATNTACTATACACACACCACTTCCC

235/310

FIGURE 232

GCCGAGCGCAAGAACCCTGCGCAGCCCAGAGCAGCTGCTGGAGGGGAATCGAGGCGCGGCTC
CGGGGATTCGGCTCGGGCCGCTGGCTCTGCTCTGCGGGGAGGGAGCGGGCCCCCGCGGGG
CCCGAGCCCTCCGGATCCGCCCCCTCCCCGGTCCCCCCCCCTCGGAGACTCCTCTGGCTGCT
CTGGGGGTTTCGCCGGGGCCGGGGACCCGCGGTCCGGGCGCCATGCGGGCATCGCTGCTGCTG
TCGGTGCTGCGGGCCCGCAGGGCCCCGTGGCCGTGGGCATCTCCCTGGGCTTCACCCCTGAGCCT
GCTCAGCGTCACCTGGGTGGAGGAGCCGTGCGGGCCAGGCCCCGCCAACCTGGAGACTCTG
AGCTGCCGCCGCGCGGCAACACCAACGCGGCGCGCCGGCCCCAACTCGGTGCAGCCCCGAGCG
GAGCGCGAGAAGCCCCGGGGCCGGGCGAAGGCGCGGGGAGAATTGGGAGCCGCGCGTCTTGCC
CTACCACCCTGCACAGCCCCGGCCAGGCCGCCAAAAAGGCCGTCAGGACCCGCTACATCAGCA
CGGAGCTGGGCATCAGGCAGAGGCTGCTGGTGGCGGTGCTGACCTCTCAGACCACGCTGCCC
ACGCTGGGCGTGGCCGTGAACCGCACGCTGGGGCACCGGCTGGAGCGTGTGGTGTTCCTGAC
GGGCGCACGGGGCCGCCGGGCCCCACCTGGCATGGCAGTGGTGACGCTGGGCGAGGAGCGAC
CCATTGGACACCTGCACCTGGCGCTGCGCCACCTGCTGGAGCAGCACGGCGACGACTTTGAC
TGGTTCTTCCTGGTGCTGACACCACCTACACCGAGGCGCACGGCCTGGCACGCCTAACTGG
CCACCTCAGCCTGGCCTCCGCCGCCACCTGTACCTGGGCGCGCCCCAGGACTTCATCGGCG
GAGAGCCCCACCCCGGCCGCTACTGCCACGGAGGCTTTGGGGTGCTGCTGTGCGCGATGCTG
CTGCAACAACCTGCGCCCCACCTGGAAGGCTGCCGCAACGACATCGTCAGTGCGCGCCCTGA
CGAGTGGCTGGGTGCTGCAATTCTCGATGCCACCGGGGTGGGCTGCACTGGTGACCACGAGG
GGGTGCACTATAGCCATCTGGAGCTGAGCCCTGGGGAGCCAGTGACAGAGGGGGACCCCTCAT
TTCCGAAGTGCCCTGACAGCCCCACCTGTGCGTGACCCCTGTGCACATGTACCAGCTGCACAA
AGCTTTCGCCCCGAGCTGAACTGGAACGCACGTACCAGGAGATCCAGGAGTTACAGTGGGAGA
TCCAGAATACCAGCCATCTGGCCGTTGATGGGGACCGGGCAGCTGCTTGGCCCCGTGGGTATT
CCAGCACCATCCCCCGGCCCTCCCGCTTTGAGGTGCTGCGCTGGGACTACTTCACGGAGCA
GCACGCTTTCTCCTGCGCCGATGGCTCACCCCGCTGCCCCACTGCGTGGGGCTGACCGGGCTG
ATGTGGCCGATGTTCTGGGGACAGCTCTAGAGGAGCTGAACCGCCGCTACCACCCGGCCTTG
CGGCTCCAGAAGCAGCAGCTGGTGAATGGCTACCGACGCTTTGATCCGGCCCCGGGTATGGA
ATACACGCTGGACTTGACAGCTGGAGGCACTGACCCCCAGGGAGGCGCGCGGCCCTCACTC
GCCGAGTGACAGCTGCTCCGGCCGCTGAGCCGCGTGGAGATCTTGCCCTGTGCCCTATGTCACT
GAGGCCTCACGTCTCACTGTGCTGCTGCCTCTAGCTGCGGCTGAGCGTGACCTGGCCCCCTGG
CTTCTTGAGGCCTTTGCCACTGCAGCACTGGAGCCTGGTGATGCTGCGGCAGCCCTGACCC
TGCTGCTACTGTATGAGCCGCGCCAGGCCAGCGCGTGGCCCATGCAGATGTCTTCGCACCT
GTCAAGGCCACGTGGCAGAGCTGGAGCGCGCTTTCCCCGGTGCCCGGGTGCCATGGCTCAG
TGTGCAGACAGCCGCACCCCTCACCACTGCGCCTCATGGATCTACTCTCCAAGAAGCACCCGC
TGGACACACTGTTCTCTGCTGGCCGGGCCAGACACGGTGCTCACGCCTGACTTCCTGAACCGC
TGCCGCATGCATGCCATCTCCGGCTGGCAGGCCTTCTTTCCCATGCATTTCCAAGCCTTCCA
CCCAGGTGTGGCCCCACCACAAGGGCCTGGGCCCCAGAGCTGGGCGGTGACACTGGCCGCT
TTGATCGCCAGGCAGCCAGCGAGGCCTGCTTCTACAACTCCGACTACGTGGCAGCCCGTGGG
CGCCTGGCGGCAGCCTCAGAACAAGAAGAGGAGCTGCTGGAGAGCCTGGATGTGTACGAGCT
GTTCTCTCACTTCTCCAGTCTGCATGTGCTGCGGGCGGTGGAGCCGGCGCTGCTGCAGCGCT
ACCGGGCCCAGACGTGCAGCGCGAGGCTCAGTGAGGACCTGTACCACCGCTGCCTCCAGAGC
GTGCTTGAGGGCCTCGGCTCCCGAACCCAGCTGGCCATGCTACTCTTTGAACAGGAGCAGGG
CAACAGCACCTGACCCCCACCTGTCCCCGTGGGCGGTGGCATGGCCACACCCACCCCACTT
CTCCCCAAAACAGAGCCACCTGCCAGCCTCGCTGGGCAGGGCTGGCCGTAGCCAGACCCC
AAGCTGGCCCACTGGTCCCCCTCTCTGGCTCTGTGGGTCCCTGGGCTCTGGACAAGCACTGGG
GGACGTGCCCCCAGAGCCACCCACTTCTCATCCAAAACCCAGTTTCCCTGCCCCCTGACGCT
GCTGATTGGGGCTGTGGCCTCCACGTATTTATGCAGTACAGTCTGCCTGACGCCAGCCCTGC
CTCTGGGCCCTGGGGGCTGGGCTGTAGAAGAGTTGTTGGGGAAGGAGGGAGCTGAGGAGGGG
GCATCTCCCAACTTCTCCCTTTTGGACCCTGCCGAAGCTCCCTGCCTTTAATAAACTGGCCA
AGTGTGGAAAAA

237/310

FIGURE 233

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELPFRGNTNAARRP
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQAACKAVRTRYISTELGIRQRLLVAVL
TSQTTLPTLGVAVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERPIGHLHLALRHLLLE
QHGDDEFDWFFLVPDTTYTEAHGLARLTGHLSLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGRCILDATGVGCTGDHEGVHYSHLELSPGEP
VQEGDPHFERSALTAHPVRDPVHMYQLHKAFARAELEPTYQEIQELQWEIQNTSHLAVDGDRA
AAWPVGIPAPSRPASRFEVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADVLGTALEELN
RRYHPALRLQKQQLVNGYRRFDPARGMEYTLDLQLEALTPQGRRPLTRRVQLLRPLSRVEI
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAALTLLLLYEPRQAQRVA
HADVFAPVKAHVAELERRFFGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTLFLLAGPDTVL
TPDFLNRCRMHAISGWQAFFPMHFQAFHPGVAPPQPGPPELGRDTGRFDRQAASEACFYNS
DYVAARGRLAAASEQEEELLES LDVYELFLHFSSLVLR AVEPALLQRYRAQTCSARLSEDL
YHRCLQSVLEGLGSRTQLAMLLFEQEQQNST

238/310

FIGURE 234

GCTCTGGCCGGCCCCGGCGATTGGTCACCGCCCGCTAGGGGACAGCCCTGGCCTCCTCTGAT
TGGCAAGCGCTGGCCACCTCCCCACACCCCTTGCGAACGCTCCCCTAGTGGAGAAAAGGAGT
AGCTATTAGCCAATTTCGGCAGGGCCCGCTTTTGTAGAAGCTTGATTTTCCTTTGAAGATGAAAG
ACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGGAACTCGGGGCGATTGGCTGGGAA
CTGTATCCACCCAAATGTCACCGATTTCTTCCTATGCAGGAAATGAGCAGACCCATCAATAA
GAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAACTGGAGGCAAAG
AGGGTTGCTCAACGCCCGCCCTCATTGGAAAACCAAATCAGATCTGGGACCTATATAGCGTG
GCGGAGGCGGGGCGATGATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCT
TTCCCCGCCCCTGAGACCCTGCAGCACCATCTGTCTATGGCGGCTGGGCTGTTTGTTTGAGC
GCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGCTCCCGGCCGCCCCGCTCCGCTGGGA
ATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGTGGCGGGAAAGCGGCCCCCAGAAC
CGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAACCTTGATGAGAAGAACCCA
GACTCCCATGGTTATGACAAGGACCCCGTTTTGGACGTCTGGAACATGCGACTTGTCTTCTT
CTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACA
GGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCTTGTGAAATACCGAGAGGCCAATGGC
CTTCCCATCATGGAATCCAACCTGCTTCGACCCCAGCAAGATCCAGCTGCCAGAGGATGAGTG
ACCAGTTGCTAAGTGGGGCTCAAGAAGCACCGCCTTCCCCACCCCCTGCCTGCCATTCTGAC
CTCTTCTCAGAGCACCTAATTAAAGGGGCTGAAAGTCTGAA

FIGURE 235

MAAGLFGLSARRLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPE
DENLYEKNPDSHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER
LVKYREANGLPIMESNCFDPSKIQLPED

e 40/310

FIGURE 236

GGCGGCTGGGCTGTTTGGTTTGAGCGCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGC
TCCCGGCCGCCCCGCTCCGCTGGGAATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCT
GTGGCGGGAAAGCGGCCCCCAGAACCGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGA
CGAAAACCTTGATGAGAAGAACCAGACTCCCATGGTTATGACAAGGACCCCGTTTTGGACG
TCTGGAACATGCGACTTGTCTTCTTCTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACC
TTTGTGGCCTATCTGCCTGACTACAGGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCT
TGTGAAATACCGAGAGGCCAATGGCCTTCCCATCATGGAATCCAAGCTTCGACCCCAGCA
AGATCCAG

241/310
FIGURE 237

GCGGCGGCTATGCGGCTTGCTCTGCTCGTCCTGTTGCTCCTGGGGCCCCGGCGGCTGGTGCCT
TGCAGAACCCCCACGCGACAGCCTGCGGGAGGAACTTGTATCAACCCGCTGCCTTCCGGGG
ACGTAGCCGCCACATTCCAGTTCCGCACGCGCTGGGATTTCGGAGCTTCAGCGGGAAGGAGTG
TCCCATTACAGGCTCTTTCCCAAAGCCCTGGGGCAGCTGATCTCCAAGTATTCTCTACGGGA
GCTGCACCTGTCATTACACAAGGCTTTTGGAGACCCGATACTGGGGGCCACCCTTCCTGC
AGGCCCCATCAGGTGCAGAGCTGTGGGTCTGGTTCCAAGACACTGTCACTGATGTGGATAAA
TCTTGGAAGGAGCTCAGTAATGTCCTCTCAGGGATCTTCTGCGCCTCTCTCAACTTCATCGA
CTCCACCAACACAGTCACTCCCACTGCCTCCTTCAAACCCCTGGGTCTGGCCAATGACACTG
ACCACTACTTTCTGCGCTATGCTGTGCTGCCGCGGGAGGTGGTCTGCACCGAAAACCTCACC
CCCTGGAAGAAGCTCTTGCCCTGTAGTTCCAAGGCAGGCCTCTCTGTGCTGCTGAAGGCAGA
TCGCTTGTTCCACACCAGCTACCACTCCAGGCAGTGCATATCCGCCCTGTTTGCAGAAATG
CACGCTGTACTAGCATCTCTGGGAGCTGAGGCAGACCCTGTCAGTTGTATTTGATGCCTTC
ATCACGGGGCAGGGAAAGAAAGACTGGTCCCTCTTCCGGATGTTCTCCGAACCCCTCACGGA
GCCCTGCCCCCTGGCTTCAGAGAGCCGAGTCTATGTGGACATCACCACTACAACCAGGACA
ACGAGACATTAGAGGTGCACCCACCCCCGACCACTACATATCAGGACGTATCCTAGGCACT
CGGAAGACCTATGCCATCTATGACTTGCTTGACACCGCCATGATCAACAACCTCTCGAAACCT
CAACATCCAGCTCAAGTGGAAGAGACCCCCAGAGAATGAGGCCCCCCCCAGTGCCCTTCCTGC
ATGCCCAGCGGTACGTGAGTGGCTATGGGCTGCAGAAGGGGGAGCTGAGCACACTGCTGTAC
AACACCCACCCATACCGGGCCTTCCCGGTGCTGCTGCTGGACACCGTACCCTGGTATCTGCG
GCTGTATGTGCACACCCTCACCATCACCTCCAAGGGCAAGGAGAACAACCAAGTTACATCC
ACTACCAGCCTGCCCAGGACCGGCTGCAACCCACCTCCTGGAGATGCTGATTGAGCTGCCG
GCCAACTCAGTCACCAAGGTTCCATCCAGTTTGAGCGGGCGCTGCTGAAGTGGACCGAGTA
CACGCCAGATCCTAACCATGGCTTCTATGTCAGCCATCTGTCTCAGCGCCCTTGTGCCCA
GCATGGTAGCAGCCAAGCCAGTGGACTGGGAAGAGAGTCCCCCTCTTCAACAGCCTGTTCCCA
GTCTCTGATGGCTCTAACTACTTTGTGCGGCTCTACACGGAGCCGCTGCTGGTGAACCTGCC
GACACCGGACTTCAGCATGCCCTACAACGTGATCTGCCTCACGTGCACTGTGGTGGCCGTGT
GCTACGGCTCCTTCTACAATCTCCTCACCCGAACCTTCCACATCGAGGAGCCCCGCACAGGT
GGCCTGGCCAAGCGGCTGGCCAACCTTATCCGGCGCGCCCGAGGTGTCCCCCACTCTGATT
CTTGCCCTTTCCAGCAGCTGCAGCTGCCGTTTCTCTCTGGGGAGGGGAGCCCAAGGGCTGTT
TCTGCCACTTGCTCTCCTCAGAGTTGGCTTTTGAACCAAAGTGCCCTGGACCAGGTCAGGGC
CTACAGCTGTGTTGTCCAGTACAGGAGCCACGAGCCAAATGTGGCATTGAAATTTGAATTAA
CTTAGAAATTCAATTCCTCACCTGTAGTGGCCACCTCTATATTGAGGTGCTCAATAAGCAAA
AGTGGTCGGTGGCTGCTGTATTGGACAGCACAGAAAAAGATTTCCATCACCAACAGAAAGGTC
GGCTGGCAGCACTGGCCAAGGTGATGGGGTGTGCTACACAGTGTATGTCACTGTGTAGTGA
TGGAGTTTACTGTTTGTGGAATAAAAAACGGCTGTTTCCGTGGAAAAA

242/310

FIGURE 238

MPLALLVLLLLGPGGWCLAEP PRDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY
RLF PKALGQLISKYSLRELHLSFTQGFWRTRYWGPPFLQAPSGAELWVWFQDTVTDV D KSWK
ELSNVLSGIFCASLNFIDSTNTVTPTASF KPLGLANDTDHYFLRYAVLPREVVCTENLTPWK
KLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCRNARCTSI SWELRQTL SVVFD AFITG
QGKKDWSLFRMFSRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT
YAIYDLLDTAMINNSRNLNIQLKWKRPPE NEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTH
PYRAFPVLLLDTPWYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANS
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAPVDWEESPLFNSLFPVSD
GSNYFVRLYTEPLL VNLPTPDFSMPYNVICLTCTVVAVCYGSFYNLLTRTFHIEEPRTGGLA
KRLANLIRRARGVPPL

243/310

FIGURE 239

CAACATGGGGTCCAGCAGCTTCTTGGTCCTCATGGTGTCTCTCGTTCTTGTGACCCTGGTGG
CTGTGGAAGGAGTTAAAGAGGGTATAGAGAAAGCAGGGGTTTGCCCAGCTGACAACGTACGC
TGCTTCAAGTCCGATCCTCCCCAGTGTACACAGACCAGGACTGTCTGGGGGAAAGGAAGTG
TTGTTACCTGCACTGTGGCTTCAAGTGTGTGATTCTGTGAAGGAACTGGAAGAAGGAGGAA
ACAAGGATGAAGATGTGTCAAGGCCATACCCTGAGCCAGGATGGGAGGCCAAGTGTCCAGGC
TCCTCCTCTACCAGGTGTCCTCAGAAATTGATGCTGGGTCTTTCTACCTCTGGGGGTCACTC
TCACTTGGCACCTGCCCCTGAGGGTCCTGAGACTTGGAATATGGAAGAAGCAATACCCAACC
CCACCAAAGAAAACCTGAGCTTGAAGTCCTTTTCCCCAAAAGAGGGAAGAGTCACAAAAAG
TCCAGACCCAGGGACGGTACTTTCCCTCTCTACCTGGTGCTCCTCCCTAATGCTCATGAAT
GGACCCCTCATGAATGAAACCAGTGGCCTTATAAGAGACCCCAAAGAGCTGCCTTGCCCTTC
TGCAATGTGTGATCACAGCTAGAAGGCACTGTCAGAGAAGAGAAACTGGTCCTCACCAGATG
CTGAATCTGCTGGTGCCTTGATCTTGGACTTCCCAGCCTCTAGAACTGTAAGAAATAAATAT
TTGCTGTTTATAATCCAA

244/310

FIGURE 240

MGSSSFVLVLMVSLVLTLLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC
YLHCGFKCVIPVKELEEGGNKDEDVSRPYPEPGWEAKCPGSSSTRCPQK

FIGURE 241

AAACTCAGCACTTGCCGGAGTGGCTCATTGTTAAGACAAAGGGTGTGCACTTCCTGGCCAGG
AAACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCAGAACCTTCC
TCTAGAACCCGACCCACCACCATGAGGTCTGCTGTGGAGATGCAGGCACCTGAGCCAAGG
CGTCCAGTGGTCTTGGCTTCTGGCTGTCTGGTCTTCTTTCTCTTCGCCCTCTTTTA
TTAAGGAGCCTCAAACAAAGCCTTCCAGGCATCAACGCACAGAGAACATTAAAGAAAGGTCT
CTACAGTCCCTGGCAAAGCCTAAGTCCCAGGCACCCACAAGGGCGAGGAGGACAACCATCTA
TGCAGAGCCAGCGCCAGAGAACAAATGCCCTCAACACACAAACCCAGCCCAAGGCCACACCA
CCGGAGACAGAGGAAAGGAGGCCAACCAGGCACCGCCGGAGGAGCAGGACAAGGTGCCCCAC
ACAGCACAGAGGGCAGCATGGAAGAGCCCAGAAAAAGAGAAAACCATGGTGAACACACTGTC
ACCCAGAGGGCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCACAATCATGGAAGAGCC
AGGACACAAAGACGACCCAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACG
GTGTGAGAGAAGCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTCCCAAAGTCA
GCACAGAATGCTGGCTCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCA
CAGCAGTCATCCACCTAAGGAGAAGAAACCTCAGGCCACCCACCCCCTGCCCCCTTCCAG
AGCCCCACGACGCAGAGAAACCAAAGACTGAAGGCCGCCAACTTCAAATCTGAGCCTCGGTG
GGATTTTGAGGAAAAATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTG
TGAAGATCAAAGCCTCCAAGTCGCTGTGGCTCCAGAACTCTTTCTGCCCAACCTCACTCTC
TTCTGGACTCCAGACACTTCAACCAGAGTGAGTGGGACCGCCTGGAACACTTTGCACCACC
CTTTGGCTTCATGGAGCTCAACTACTCCTTGGTGCAGAAGGTGCTGACACGCTTCCCTCCAG
TGCCCCAGCAGCAGCTGCTCCTGGCCAGCCTCCCCGCTGGGAGCCTCCGGTGATCACCTGT
GCCGTGGTGGGCAACGGGGGCATCCTGAACAACTCCACATGGGCCAGGAGATAGACAGTCA
CGACTACGTGTTCCGATTGAGCGGAGCTCTCATTAAAGGCTACGAACAGGATGTGGGGACTC
GGACATCCTTCTACGGCTTTACCGCCTTCTCCCTGACCCAGTCACTCCTTATATTGGGCAAT
CGGGGTTTCAAGAACGTGCCTCTTGGGAAGGACGTCCGCTACTTGCACTTCTGGAAGGCAC
CCGGGACTATGAGTGGCTGGAAGCACTGCTTATGAATCAGACGGTGATGTCAAAAAACCTTT
TCTGGTTCAGGCACAGACCCCAGGAAGCTTTTCGGGAAGCCCTGCACATGGACAGGTACCTG
TTGCTGCACCCAGACTTTCTCCGATACATGAAGAACAGGTTTCTGAGGTCTAAGACCCTGGA
TGGTGCCCACTGGAGGATATACCGCCCCACCACTGGGGCCCTCCTGCTGCTCACTGCCCCTC
AGCTCTGTGACCAGGTGAGTGCTTATGGCTTCATCACTGAGGGCCATGAGCGCTTTTCTGAT
CACTACTATGATACATCATGGAAGCGGCTGATCTTTTACATAAACCATGACTTCAAGCTGGA
GAGAGAAGTCTGGAAGCGGCTACACGATGAAGGGATAATCCGGCTGTACCAGCGTCCTGGTC
CCGGAAGTGCCAAAGCCAAGAACTGAACCGGGGCCAGGGCTGCCATGGTCTCCTTGCCCTGCTC
CAAGGCACAGGATACAGTGGGAATCTTGAGACTCTTTGGCCATTTCCCATGGCTCAGACTAA
GCTCCAAGCCCTTCAGGAGTTCAGGGAACACTTGAACCATGGACAAGACTCTCTCAAGAT
GGCAAATGGCTAATTGAGGTCTGAAGTTCTTCAGTACATTGCTGTAGGTCTGAGGCCAGG
GATTTTTTAATTAAATGGGGTGATGGGTGGCCAATACCACAATTCTGCTGAAAAACACTCTT
CCAGTCCAAAAGCTTCTTGATACAGAAAAAGAGCCTGGATTTACAGAAACATATAGATCTG
GTTTGAATTCCAGATCGAGTTACAGTTGTGAAATCTTGAAGGTATTACTTAACTTCACTAC
AGATTGTCTAGAAGACCTTTCTAGGAGTTATCTGATTCTAGAAGGGTCTATACTTGTCTTG
TCTTTAAGCTATTTGACAACTCTACGTGTTGTAGAAAACTGATAATAATAACAAATGATTGTT
GTCCATGGAAAGGCAAATAAATTTTCTACAGTGAAAAA

245/310

FIGURE 242

MRSCLWRCRHLSQGVQWSLLLAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP
KSQAPTRARRTTIYAEPAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW
KSPEKEKTMVNTLSPRGQDAGMASGRTEAQSWKSQDTKTTQGNGGQTRKLTASRTVSEKHQG
KAATTAKTLIPKSQHRMLAPTGA VSTRTRQKGVTTAVIPPEKKPQATPPPAPFQSPTTQRN
QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLPNLTLFLDSRHF
NQSEWDRLEHFAPPPFGFMELNYSLVQKVTRFPFPVQQQLLLASLPAGSLRCITCAVVGNGG
ILNNSHMGQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGF TAFSLTQSLLILGNRGFKNVP
LGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFL
RYMKNRFLRSKTLDGAWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYDTSW
KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

247/310

FIGURE 243

CGATGCGCGGACCCGGGCACCCCCTCCTCCTGGGGCTGCTGCTGGTGCTGGGGCCTTCGCCG
GAGCAGCGAGTGGAATTGTTCCCTCGAGATCTGAGGATGAAGGACAAGTTTCTAAAACACCT
TACAGGCCCTCTTTATTTTAGTCCAAAGTGCAGCAAACACTTCCATAGACTTTATCACAACA
CCAGAGACTGCACCATTCTGCATACTATAAAAGATGCGCCAGGCTTCTTACCCGGCTGGCT
GTCAGTCCAGTGTGCATGGAGGATAAGTGAGCAGACCGTACAGGAGCAGCACACCAGGAGCC
ATGAGAAGTGCCTTGGAACCAACAGGGAAACAGAACTATCTTTATACACATCCCCTCATGG
ACAAGAGATTTATTTTTCGAGACAGACTCTTCCATAAGTCCTTTGAGTTTGTATGTTGTTG
ACAGTTTGCAGATATATATTCGATAAATCAGTGTACTTGACAGTGTTATCTGTCACTTATTT

248/310

FIGURE 244

MRGPGHPLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFHRLYHNT
RDCTIPAYYKRCARLLTRLAVSPVCMEDK

FIGURE 245

GGGCTGGGCCCCGCGCGAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG
CCCGACCCCGGCGCGCGCCAGCCCCACCAATGCCACCCGCGGGGCTCCGCCGGGCGCGCGCG
CTCACCGCAATCGCTCTGTTGGTGCTGGGGGCTCCCCTGGTGCTGGCCGGCGAGGACTGCCT
GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT
GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCGAGAGG
CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT
CCTCTTTGTTGCTGTGGTTGCCACCACCATCTGCTGCTTCCTCTGTTCCCTGTTGCTACCTGT
ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTGAAGGCCAGGAGATTCCAATGACAGGCATC
CCAGTGCAGCCAGTATACCCATACCCCCAGGACCCCAAAGCTGGCCCTGCACCCCCACAGCC
TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCCAATATCCACTCTACCCAGCTGGGCCCC
CAGTCTACAACCCTGCAGCTCCTCCTCCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC
TGAGGAACCAGCCATGTCTCTGCTGCCCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT
CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA
GCCAAGCCCTGGGCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA
ACTATGAGGGGTTGGGGGGAGGGCTTGGAATTATGGGCTATTTTTACTGGGGGCAAGGGAGG
GAGATGACAGCCTGGGTACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCG
CCAGGAAGGCTGGGGCCCTACTGTTTGTCCCCTCTGGGCTGGGGTGGGGGGAGGGAGGAGGT
TCCGTCAGCAGCTGGCAGTAGCCCTCCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG
CTAGATTAAAGCTGTAAAGACAAAA

FIGURE 246

MPPAGLRRAAPLTAIALLVLGAPLVLAGEDCLWYLDRNGSWHPGFNCEFFTCCGTCTYHRYC
CRDLTLLITERQQKHCLAFSPKTIAGIASAVILFVAVVATTICCFLCSCCYLYRRRQQLQSP
FEGQEIPMTGIPVQPVYPYPQDPKAGPAPPQPGFMYPPSGPAPQYPLYPAGPPVYNPAAPPP
YMPPQPSYPGA